

> 0 <  
0|/0 IncelliGenetics  
> 0 <

GENALIGN - Multiple Sequence Alignment Program  
Release 5.4

Thu 4 Aug 105 5:43:31- PST

Solution Parameters:

Amino Alphabet = Identity  
Output line length = 80  
Compress = Off  
Histogram = Off  
Randomization = Off  
AMINO-Res-length = 2  
Deletion-weight = 5.00  
Length-factor = 0  
Matching-weight = 1.00  
NUCLEIC-Res-length = 4  
Spread-factor = 50

Clustured order of selected sequences:

3. US-10-634-108-5 (1-277)
1. US-10-634-108-2 (1-245)
2. US-10-634-108-4 (1-243)

Region Alignment: (listed in Clustured order)

```
US-10-634- 1 mmpaaelgqclseraglcrlcllcaaglphtmhPQGRASPQILLGLFLVLLLLQLSAP
US-10-634- 1 MHPQGRASPQILLGLFLVLLLLQLSAP
US-10-634- 1 MHPQGRASPQILLGLFLVLLLLQLSAP
US-10-634- 1 MHPQGRASPQILLGLFLVLLLLQLSAP
consensus mmpaaelgqclseraglcrlcllcaaglphtmhPQGRASPQILLGLFLVLLLLQLSAP

US-10-634- 62 SSASENPVKVKQKALIRQREVVDLYNGMCLQGPAGVGRDSSPGANGIPGTPGIPGRDGFKG
US-10-634- 30 SSASENPVKVKQKALIRQREVVDLYNGMCLQGPAGVGRDSSPGANGIPGTPGIPGRDGFKG
US-10-634- 28 SSASENPVKVKQKALIRQREVVDLYNGMCLQGPAGVGRDSSPGANGIPGTPGIPGRDGFKG
consensus SSASENPVKVKQKALIRQREVVDLYNGMCLQGPAGVGRDSSPGANGIPGTPGIPGRDGFKG

US-10-634- 123 EKGECLERESFEESWTPNTKQCSWSLSNYGIDLGKIAECTFTKMRNSALRVLPFSGSLRLKC
US-10-634- 91 EKGECLERESFEESWTPNTKQCSWSLSNYGIDLGKIAECTFTKMRNSALRVLPFSGSLRLKC
US-10-634- 89 EKGECLERESFEESWTPNTKQCSWSLSNYGIDLGKIAECTFTKMRNSALRVLPFSGSLRLKC
consensus EKGECLERESFEESWTPNTKQCSWSLSNYGIDLGKIAECTFTKMRNSALRVLPFSGSLRLKC

US-10-634- 184 RNACCQRRWYFTFNGAECGSLPIEAIIVLDQSPELNSTINIHRTSSVEGLCEGIGAGLVLD
US-10-634- 152 RNACCQRRWYFTFNGAECGSLPIEAIIVLDQSPELNSTINIHRTSSVEGLCEGIGAGLVLD
US-10-634- 150 RNACCQRRWYFTFNGAECGSLPIEAIIVLDQSPENNSTINIHRTSSVEGLCEGIGAGLVLD
consensus RNACCQRRWYFTFNGAECGSLPIEAIIVLDQSPENNSTINIHRTSSVEGLCEGIGAGLVLD

US-10-634- 245 VAIWVGTCSDYPKGDASTGMNSVSRRIIEELPK
US-10-634- 213 VAIWVGTCSDYPKGDASTGMNSVSRRIIEELPK
```

US-10-634- 211 VAIWVGTCSDYPKGDASTGMNSVSRRIIEELPK  
consensus VAIWVGTCSDYPKGDASTGMNSVSRRIIEELPK

Alignment score = 147.00

Scoring matrix:

	1	2	3
1		222	85
2			63
3			

-----

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:04:29 ; Search time 93.7059 Seconds  
(without alignments)  
1002.955 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303

Sequence: 1 MRPGPAPASQRLRGILL.....GDASTGMSVSRITIELPK 243

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: geneseqp1980s:\*  
2: geneseqp1980s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	AAB36667	Aab36667 Human sec
2	1303	100.0	243	ABG96340	Abg96340 Human ova
3	1303	100.0	243	ABJ05554	Abj05554 Breast ca
4	1303	100.0	243	ABB80978	Abb80978 Human REM
5	1303	100.0	243	ABR58546	AbR58546 Human can
6	1303	100.0	243	ABR48227	AbR48227 Human bla
7	1303	100.0	243	ABG75758	AbG75758 Human REM
8	1303	100.0	243	ABU56607	AbU56607 Lung canc
9	1303	100.0	243	ADB80510	AdB80510 Ovarian c
10	1303	100.0	243	ADN39855	Adn39855 Cancer/an
11	1303	100.0	243	ADN38732	Adn38732 Cancer/an
12	1303	100.0	243	ADL70254	Adl70254 IBFL301 p
13	1303	100.0	278	AAB08856	Aab08856 Amhu aci
14	1303	100.0	278	AAM25746	Aam25746 Human pro
15	1303	100.0	278	ABG96338	Abg96338 Human ova
16	1303	100.0	278	ABR47627	AbR47627 Breast ca
17	1300	99.8	243	AAV91529	Aay91529 Human sec
18	1300	99.8	278	ADL71601	Adl71601 Novel hum
19	1300	99.8	243	ABP66631	Abp66631 Human pan
20	1298	99.6	243	ABJ37031	Abj37031 Human bre
21	1294	99.3	243	AAV99462	Aay99462 Human PRO
22	1294	99.3	243	AAV99441	Aay99441 Human PRO
23	1294	99.3	243	AAB66190	Aab66190 Protein o
24	1294	99.3	243	AAB66211	Aab66211 Protein o
25	1294	99.3	243	AAU29206	Aau29206 Human PRO

26	1294	99.3	243	4	AAB87586	Aab87586 Human PRO
27	1294	99.3	243	5	ABG95911	Abg95911 Human sec
28	1294	99.3	243	5	ABB84939	Abb84939 Human PRO
29	1294	99.3	243	5	AAE20462	Aae20462 Human tum
30	1294	99.3	243	5	ABR95545	Abbr95545 Human ang
31	1294	99.3	243	5	ABP66636	Abp66636 Human pan
32	1294	99.3	243	5	ABG78938	Abg78938 Human bre
33	1294	99.3	243	5	ABU58582	AbU58582 Human PRO
34	1294	99.3	243	6	ABU88130	Abu88130 Novel hum
35	1294	99.3	243	6	ABU84445	Abu84445 Human sec
36	1294	99.3	243	6	ABR66319	AbR66319 Human sec
37	1294	99.3	243	6	ABR65709	AbR65709 Human sec
38	1294	99.3	243	6	ABU96649	AbU96649 Human sec
39	1294	99.3	243	6	ABJ37761	Abj37761 Human tum
40	1294	99.3	243	6	ABU82888	Abu82888 Human PRO
41	1294	99.3	243	6	ABU90009	Abu90009 Novel hum
42	1294	99.3	243	6	ABR68258	AbR68258 Human sec
43	1294	99.3	243	6	ABU96311	AbU96311 Novel hum
44	1294	99.3	243	6	ABU92742	Abu92742 Human sec
45	1294	99.3	243	6	ABO08819	AbO08819 Human sec

## ALIGNMENTS

RESULT 1  
AAB36667 standard; protein; 243 AA.  
ID AAB36667;  
AC AAB36667;  
XX  
XX  
DT 14-MAR-2001 (first entry)  
XX  
DE Human secretory protein TGC-628 SEQ ID NO:7.  
XX  
XX  
KW Human; secretory protein; cancer; immune disease; infectious disease;  
KW lung function disorder; liver function disorder; antiinflammatory;  
KW gastrointestinal disorder; cytostatic; hematopoietic; anticosgulant;  
KW immunomodulatory; hepatotropic; cell proliferation-stimulant;  
KW cell migratory agent; cell differentiation-inducer.  
XX  
OS Homo sapiens.  
XX  
PN W0200071581-A1.  
XX  
PD 30-NOV-2000.  
XX  
PF 19-MAY-2000; 2000WO-JP003221.  
XX  
PR 20-MAY-1999; 99JP-00140229.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Itoh Y, Mogi S, Tanaka H, Ohkubo S, Ogi K;  
XX  
PI WPI; 2001-032023/04.  
XX  
DR N-PSDB; AAC90707.  
XX

Novel secretory protein and its salt with e.g. anti-cancer, anti-inflammatory and hematopoietic, effects, applicable as drugs in remedies and preventives to treat diseases like cancer and immune diseases.  
Claim 1, Page 89-90; 122pp, Japanese.  
AAC90701 to AAC90715 encode the human secretory proteins given in AAB36661 to AAB36675. The proteins can have cytostatic, anti-inflammatory, haematopoietic, anti-coagulant, immunomodulatory and hepatotropic activities, and can be used as cell migratory agents, cell proliferation-stimulants and cell differentiation-inducers. The proteins are useful in the treatment and prevention of diseases such as cancer, lung function disorder, liver function disorder, gastrointestinal disorder and immune diseases. AAC90716 to AAC90755 represent PCR primers which are used in the exemplification of the present invention

```
XX
SQ Sequence 243 AA;
Query Match 100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLRGLLLLLQLPAPSSASIPKQKQAKQREVDVLYNMGICQGA 60
   |||
DB 1 MRPGPASPQRRLRGLLLLLQLPAPSSASIPKQKQAKQREVDVLYNMGICQGA 60
   |||
QY 61 GVPGRDGSFGANGIPGTPIGPDGFKGEGKCECLRESFEESWTPTYKQCSWSSLNYGIDL 120
   |||
DB 61 GVPGRDGSFGANGIPGTPIGPDGFKGEGKCECLRESFEESWTPTYKQCSWSSLNYGIDL 120
   |||
QY 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGSLPIEAIITYLQ 180
   |||
DB 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGSLPIEAIITYLQ 180
   |||
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
   |||
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
   |||
QY 241 LPK 243
   |||
DB 241 LPK 243
   |||

RESULT 2
ABG96340 ID ABG96340 standard; protein; 243 AA.
XX
AC ABG96340;
XX
XX 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M450.
XX
XX Human ovarian cancer; marker; cancer; familial history; brain disorder;
   central nervous system disorder; bacterial meningitis; viral meningitis;
   Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
   brain herniation; inflammation; encephalitis; testicular disorder;
   nondercutaneous granulomatous orchitis; connective tissue disorder;
   heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
   histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
PN WO200271928-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002WO-US007826.
XX
PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Monahan JE, Gannavarapu M, Hoersch S, Kamatkar S, Kovatis SG,
PI Meyers RE, Morrissey AP, Oland PJ, Sen A, Vleby PO, Mills GB,
PI Baat RC, Lu K, Schmandt RE, Zhao X, Giatt K;
XX
XX MPI: 2002-723277/78.
XX
DR N-PSDB; AB876436.
XX
PT Assessing whether a patient is afflicted with ovarian cancer, useful in
   assessing the stage or progression of the disease, comprises comparing
```

```
PT the expression level of a cancer marker in a sample from a patient and
   from a non cancer patient.
XX
PS Disclosure; Page 252; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
   patient is afflicted with ovarian cancer. The method involves comparing
   the expression level of a marker in a patient sample and the normal level
   of expression of the marker in a control non-ovarian cancer sample, where
   the marker is selected from 363 cancer markers described in the
   specification. The method of the invention is useful in diagnosing or
   characterizing cancer, in detecting the presence of cancer as early as
   possible, and the recurrence of ovarian cancer. The method may also be of
   particular use with patients having an enhanced risk of developing
   ovarian cancer (e.g. patients having a familial history of ovarian
   cancer). The cancer markers may be used in the management and treatment
   of e.g. brain and central nervous system disorders (e.g. bacterial and
   viral meningitis, Alzheimer's disease or Parkinson's disease), brain
   disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
   inflammations (e.g. bacterial or viral meningitis or encephalitis),
   testicular disorders (e.g. nondercutaneous granulomatous orchitis),
   connective tissue disorders, or heart disorders (e.g. ischaemic heart
   disease or atherosclerosis). The compositions and methods may also be
   used in assessing the histological type of neoplasm associated with
   ovarian cancer, monitoring the progression of ovarian cancer, determining
   whether ovarian cancer has metastasized or is likely to metastasize,
   selecting a composition for inhibiting ovarian cancer, assessing the
   ovarian carcinogenic potential of a compound, or inhibiting ovarian
   cancer or at risk of developing ovarian cancer. The present amino acid
   sequence represents one of the ovarian cancer markers described in the
   invention
XX
SQ Sequence 243 AA;
Query Match 100.0%; Score 1303; DB 5; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLRGLLLLLQLPAPSSASIPKQKQAKQREVDVLYNMGICQGA 60
   |||
DB 1 MRPGPASPQRRLRGLLLLLQLPAPSSASIPKQKQAKQREVDVLYNMGICQGA 60
   |||
QY 61 GVPGRDGSFGANGIPGTPIGPDGFKGEGKCECLRESFEESWTPTYKQCSWSSLNYGIDL 120
   |||
DB 61 GVPGRDGSFGANGIPGTPIGPDGFKGEGKCECLRESFEESWTPTYKQCSWSSLNYGIDL 120
   |||
QY 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGSLPIEAIITYLQ 180
   |||
DB 121 GKIAECTFTKMSNSALRYLFSGSLRLKCRNACCQRMWYTFNGAECGSLPIEAIITYLQ 180
   |||
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
   |||
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
   |||
QY 241 LPK 243
   |||
DB 241 LPK 243
   |||

RESULT 3
ABJ05554 ID ABJ05554 standard; protein; 243 AA.
XX
AC ABJ05554;
XX
XX 14-NOV-2002 (first entry)
XX
DE Breast cancer-associated protein 19.
XX
XX Breast cancer; breast cancer-associated gene sequence; drug development;
   pharmacogenetics; biosensor development.
XX
XX Unidentified.
```



XX MO200259377-A2.  
 FN 01-AUG-2002.  
 XX 24-JAN-2002; 2002MO-US002242.  
 XX 24-JAN-2001; 2001US-0263965P.  
 PR 02-FEB-2001; 2001US-0265928P.  
 PR 09-APR-2001; 2001US-00829472.  
 PR 04-MAY-2001; 2001US-0282698P.  
 PR 29-MAY-2001; 2001US-0288590P.  
 XX (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX Mack DH, Gish KC, Afar D;  
 PI WPI; 2002-583738/62.  
 DR N-PSDB; ABT07711.  
 XX Detecting a breast cancer-associated transcript in a patient's cell,  
 PT useful for diagnosing breast cancer, comprises contacting a biological  
 PT sample with a polynucleotide that selectively hybridizes with breast  
 PT cancer nucleic acids.  
 XX  
 XX Disclosure; Page 364; 414pp; English.  
 PS The invention comprises a method of detecting a breast cancer-associated  
 CC transcript in a cell from a patient. The method of the invention involves  
 CC contacting a biological sample from the patient with a nucleotide that  
 CC hybridizes to one of the 69 breast cancer-associated gene sequences shown  
 CC in the specification. The method of the invention is useful in the  
 CC diagnosis or prognosis of breast cancer, and for detecting genes that are  
 CC up or down-regulated in breast cancer cells. Genes identified by the  
 CC method of the invention can be used in diagnostic purposes and also as  
 CC targets for screening for therapeutic compounds that modulate breast  
 CC cancer (e.g. hormones or antibodies). Identification of genes that are  
 CC over or under expressed in breast cancer can additionally provide high-  
 CC resolution, high-sensitivity datasets which can be used in the areas of  
 CC diagnostics, therapeutics, drug development, pharmacogenetics, protein  
 CC structure and biosensor development. Amino acid sequences AB05536 -  
 CC AB05604 represent the proteins encoded by the 69 breast cancer-  
 CC associated genes of the invention  
 XX  
 XX Sequence 243 AA;  
 SQ  
 Query Match 100.0%; Score 1303; DB 5; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9e-121;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRDQGPASPORLRLGILLILLQLPAPSSASEIPKQKQALRQREVVLDLYNGMCLQGPA 60  
 DB 1 MRDQGPASPORLRLGILLILLQLPAPSSASEIPKQKQALRQREVVLDLYNGMCLQGPA 60  
 QY 61 GVGGRGSSPRANGIPGPIGRDGFGEKGECLRESESSSWPNYKQCSWSSINVGIDL 120  
 DB 61 GVGGRGSSPRANGIPGPIGRDGFGEKGECLRESESSSWPNYKQCSWSSINVGIDL 120  
 QY 121 GKIAECTFTMRNSALRVLFSGSLRLKCRNACQQRVYFTFNGAECGSPPIFAITYLDQ 180  
 DB 121 GKIAECTFTMRNSALRVLFSGSLRLKCRNACQQRVYFTFNGAECGSPPIFAITYLDQ 180  
 QY 181 GSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNSVSRITIER 240  
 DB 181 GSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNSVSRITIER 240  
 QY 241 LPK 243  
 DB 241 LPK 243

ABB80978  
 ID ABB80978 standard; protein; 243 AA.  
 XX  
 AC ABB80978;  
 XX  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Human REMODELIN polypeptide.  
 XX  
 XX REMODELIN; vulnerrary; vasotrophic; cyostatic; osteopathic; collagen;  
 KW gene therapy; bone; human.  
 KM  
 OS Homo sapiens.  
 XX  
 XX WO200242487-A2.  
 XX  
 XX 30-MAY-2002.  
 XX  
 XX 19-OCT-2001; 2001MO-US050940.  
 XX  
 XX 19-OCT-2000; 2000US-00692081.  
 XX  
 XX (MAIN-) MAINE MEDICAL CENT RES INST.  
 XX  
 XX Linder V, Friesel R;  
 PI  
 XX WPI; 2002-590472/63.  
 DR N-PSDB; ABB86481.  
 XX  
 XX New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
 PT negative remodeling, fibrosis, collagen disease, and bone disease in a  
 PT mammal.  
 XX  
 PS Claim 5; Fig 4B; 187pp; English.  
 XX  
 XX The invention relates to mammalian REMODELIN polypeptides and encoding  
 CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
 CC negative remodeling or fibrosis, bone disease such as osteogenesis  
 CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
 CC bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are  
 CC used to affect cellular gene expression in a mammal, where the cellular  
 CC gene is from transforming growth factor (TGF)-beta1, collagen I(alpha1),  
 CC osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein  
 CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
 CC comprising antisense REMODELIN sequences are useful for treating diseases  
 CC mediated by abnormal expression of a REMODELIN molecule in a human such  
 CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or  
 CC hypertrophic scar formation. REMODELIN is useful in the development of  
 CC bone during mammalian embryogenesis, proliferation and/or migration.  
 CC cellular signaling, adventitial fibrosis, negative remodeling and arterial  
 CC restenosis, smooth muscle cell proliferation, and arterial remodeling and  
 CC useful in gene therapy. REMODELIN is useful for elucidating the function  
 CC of REMODELIN molecules in a cell, to identify a compound that affects  
 CC REMODELIN expression and/or TGF-beta signaling, as a potential  
 CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
 CC formation, to promote bone fracture healing, and to increase or decrease  
 CC expression of REMODELIN in mammal. It is also useful for producing  
 CC recombinant cell and transgenic non-human mammals which are useful tools  
 CC for the study of REMODELIN action, for identifying novel diagnostics and  
 CC therapeutics for treatment, and for elucidating the cellular roles of  
 CC REMODELIN. The present sequence represents a human REMODELIN polypeptide  
 XX  
 XX Sequence 243 AA;  
 SQ  
 Query Match 100.0%; Score 1303; DB 5; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9e-121;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRDQGPASPORLRLGILLILLQLPAPSSASEIPKQKQALRQREVVLDLYNGMCLQGPA 60  
 DB 1 MRDQGPASPORLRLGILLILLQLPAPSSASEIPKQKQALRQREVVLDLYNGMCLQGPA 60

```

QY 61 GVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKQCSMSLNYGIDL 120
    |||||
Db 61 GVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKQCSMSLNYGIDL 120
QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180
    |||||
Db 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180
QY 181 GSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRITIEE 240
    |||||
Db 181 GSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRITIEE 240
QY 241 LPK 243
    |||
Db 241 LPK 243

RESULT 5
ABR58546
ID ABR58546 strand; protein; 243 AA.
XX
AC ABR58546;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human cancer related protein SEQ ID NO:203.
XX
KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KW heart disease; atherosclerosis; endometritis.
XX
OS Homo sapiens.
XX
PN MO2003025138-A2.
XX
PD 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-US029560.
XX
PR 17-SEP-2001; 2001US-0323469P.
PR 20-SEP-2001; 2001US-0323887P.
PR 13-NOV-2001; 2001US-0350666P.
PR 08-FEB-2002; 2002US-0355145P.
PR 08-FEB-2002; 2002US-0355257P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
PI Zlotnik A;
XX
DR N-PSDB; ACC72666.
XX
WPI: 2003-354600/33.
XX
DR N-PSDB; ACC72666.

New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
PS Claim 12, Page 736; 767pp; English.
XX
XX
CC The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-
CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer in
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for

```

```

CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometritis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX
SO Sequence 243 AA;
XX
Query Match 100.0%; Score 1303; DB 6; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPORLRGLLILLQLPAPSSASRIPKGKQKAOIROREVDLYNMGCLQGPA 60
Db 1 MRPGPASPORLRGLLILLQLPAPSSASRIPKGKQKAOIROREVDLYNMGCLQGPA 60
QY 61 GVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKQCSMSLNYGIDL 120
    |||||
Db 61 GVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESFEESWTPNYKQCSMSLNYGIDL 120
QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180
    |||||
Db 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCQRMWFTFNGACSGPLPIEAIYYLDQ 180
QY 181 GSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRITIEE 240
    |||||
Db 181 GSEPMNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRITIEE 240
QY 241 LPK 243
    |||
Db 241 LPK 243

RESULT 6
ABR48227
ID ABR48227 strand; protein; 243 AA.
XX
AC ABR48227;
XX
DT 12-JUN-2003 (first entry)
XX
DE Human bladder cancer associated protein sequence SEQ ID NO:175.
XX
KW Human; bladder cancer; cytostatic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
PN MO2003003906-A2.
XX
PD 16-JAN-2003.
XX
PF 03-JUL-2002; 2002WO-US021338.
XX
PR 03-JUL-2001; 2001US-0302814P.
PR 03-AUG-2001; 2001US-0310099P.
PR 08-NOV-2001; 2001US-0343705P.
PR 13-NOV-2001; 2001US-0350666P.
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
PI Mack DH, Aziz N;
XX
DR N-PSDB; ACC51043.
XX
WPI: 2003-201532/19.
XX
DR N-PSDB; ACC51043.

Detecting a bladder cancer-associated transcript in a cell from a
PT patient, comprises contacting a biological sample from the patient with a
PT bladder cancer-associated polynucleotide or antibody.
XX
PS Claim 10; Page 289-290; 307pp; English.
XX
XX

```

CC The present invention describes a method for detecting a bladder cancer-associated transcript in a cell from a patient. The method comprises contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence that is 80 % identical to a table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059 encode the human bladder cancer-associated proteins given in ABR48146 to ABR48242). Bladder cancer-associated sequences from the present invention have cytostatic activities, and can be used in antisense gene therapy and in vaccine production. The method can be used for detecting a bladder cancer-associated transcript in a cell from a patient. The method is useful in diagnosing or treating bladder cancer and in screening for compounds that modulate bladder cancer, such as hormones or antibodies. The nucleic acid molecules from the present invention may be used in various screening and diagnostic methods, and for gene therapy, vaccine and/or antisense/inhibition applications

CC Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 6; Length 243;  
Best Local Similarity 100.0%; Pred. No. 9e-121; Mismatches 0; Gaps 0;  
Matches 243; Conservative 0; Indels 0; Gaps 0;

QY 1 MRPOGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCLQGPA 60  
1 MRPOGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCLQGPA 60  
Db 61 GVPGRDGSFGANGIPETPGIPGRDGFKEGECLEBSFESWTPTYKQCSWSLNYGIDL 120  
61 GVPGRDGSFGANGIPETPGIPGRDGFKEGECLEBSFESWTPTYKQCSWSLNYGIDL 120  
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
Db 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPRGDASTGMNSVSRITIEE 240  
181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPRGDASTGMNSVSRITIEE 240  
Db 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPRGDASTGMNSVSRITIEE 240  
QY 241 LPK 243  
241 LPK 243  
Db 241 LPK 243

#### RESULT 7

ABG75758  
ID ABG75758 standard; protein; 243 AA.

XX ABG75758;

DT 28-APR-2003 (first entry)

XX Human REMODELIN protein.

XX Human; adventitia-inducible; REMODELIN; REMODEL;  
KM adventitia induced bone expressed molecule; AIBE; TGF-beta;  
KM transforming growth factor beta; adventitia; vascular remodelling;  
KM restenosis; vascular injury; antisense therapy; TGF-beta signalling;  
KM TGF-beta receptor type II; arterial remodelling; bone formation;  
KM cartilage formation; osteogenesis imperfecta; Bethlem myopathy;  
KM dystrophic epidermolysis bullosa; negative remodelling; wound healing;  
KM arterial stenosis; fibrosis; calcification; transplant;  
KM heart valve transplant; osteopathic; antiarteriosclerotic; vulnerary.

OS Homo sapiens.

XX US2002161211-A1.

PD 31-OCT-2002.

PF 19-OCT-2001; 2001US-00045992.

XX 19-OCT-2000; 2000US-00692081.

XX

PA (LIND/) LINDNER V.  
PA (FRIE/) FRIESEL R E.  
XX Linde V, Friesel R E;  
PI Linde V, Friesel R E;  
XX WPI: 2003-238238/23.  
DR N-PSDB, ABX11341.

PT New isolated REMODELIN nucleic acid and polypeptide, useful for mediating arterial remodelling, formation of bone and cartilage, and the diagnosis and treatment of disorders associated with aberrant expression of REMODELIN.

PS Claim 3; Fig 4B; 81pp; English.

CC The invention discloses an isolated nucleic acid encoding a mammalian CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL CC and/or adventitia induced bone expressed molecule, AIBE). REMODELIN CC expression was induced by transforming growth factor beta (TGF-beta), CC which is important because proliferative events occurring in the CC adventitia contribute to vascular remodelling and restenosis in response CC to vascular injury and TGF-beta has been shown to be a factor involved in CC this. Also disclosed is an antibody raised against REMODELIN and methods CC for treating a disease mediated by abnormal expression of a REMODELIN in CC a human, for identifying a compound that affects or reduces expression of REMODELIN in a cell (e.g. antisense therapy), for identifying a compound CC that affects TGF-beta signalling and for increasing or reducing REMODELIN CC expression in a mammal, comprising administering a REMODELIN expression CC increasing or reducing TGF-beta to the mammal, thereby increasing CC REMODELIN expression or inhibiting signalling via the TGF-beta receptor CC type II and reducing expression of REMODELIN in the mammal. The methods CC and compositions of the present invention are useful for mediating CC arterial remodelling, formation of bone and cartilage and the diagnosis CC and treatment of disorders associated with aberrant expression of REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis CC bullosa, Bethlem myopathy, negative remodelling, wound healing, arterial CC stenosis, vessel injury, fibrosis and calcification of a transplant, CC preferably a heart valve transplant. The sequence presented is the human REMODELIN protein

XX Sequence 243 AA;

Query Match 100.0%; Score 1303; DB 6; Length 243;  
Best Local Similarity 100.0%; Pred. No. 9e-121; Mismatches 0; Gaps 0;  
Matches 243; Conservative 0; Indels 0; Gaps 0;

QY 1 MRPOGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCLQGPA 60  
1 MRPOGPAPSPORLRLGLLLLLQLPAPSSASSEIPKQKQAKLRQREVVDLYNGMCLQGPA 60  
Db 61 GVPGRDGSFGANGIPETPGIPGRDGFKEGECLEBSFESWTPTYKQCSWSLNYGIDL 120  
61 GVPGRDGSFGANGIPETPGIPGRDGFKEGECLEBSFESWTPTYKQCSWSLNYGIDL 120  
QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
Db 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQWRWTFNGACSGPLPIEAIITYLQD 180  
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPRGDASTGMNSVSRITIEE 240  
181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPRGDASTGMNSVSRITIEE 240  
Db 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPRGDASTGMNSVSRITIEE 240  
QY 241 LPK 243  
241 LPK 243  
Db 241 LPK 243

#### RESULT 8

ABU56607  
ID ABU56607 standard; protein; 243 AA.

XX ABU56607;

XX

```

DT 02-APR-2003 (first entry)
XX Lung cancer-associated polypeptide #200.
DB Lung cancer-associated polypeptide #200.
XX Lung cancer-associated polypeptide; cytostatic; emphysema;
XX antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
XX
OS Unidentified.
XX
PN WO200286443-A2.
XX
PD 31-OCT-2002.
XX
PF 18-APR-2002; 2002WO-US012476.
XX
PR 18-APR-2001; 2001US-0284770P.
XX
PR 10-MAY-2001; 2001US-0290492P.
XX
PR 09-NOV-2001; 2001US-0339245P.
XX
PR 13-NOV-2001; 2001US-0350666P.
XX
PR 29-NOV-2001; 2001US-0334370P.
XX
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
PI Aziz N, Murray R;
XX
DR WPI; 2003-093161/08.
XX
DR N-PSDB; ABX76336.
XX
PT Detecting a lung cancer-associated transcript in a cell from a patient
PT for treating lung cancer, by contacting a biological sample from the
PT patient with a polynucleotide that exhibits increased or decreased
PT expression in lung cancer.
XX
PS Claim 27; Page 337; 453pp; English.
XX
CC The invention relates to a method for detecting a lung cancer-associated
CC transcript in a cell from a patient, comprising contacting a biological
CC sample from the patient with a polynucleotide that selectively hybridizes
CC to a sequence that is at least 80 % identical to a gene that exhibits
CC increased or decreased expression in lung cancer samples. Lung cancer-
CC associated polynucleotides and polypeptides are used for identifying a
CC compound that modulates a lung cancer-associated polypeptide, for
CC inhibiting proliferation of a lung cancer-associated cell to treat lung
CC cancer in a patient and for treating a mammal having lung cancer by
CC administering a modulatory compound identified. The methods are useful
CC for treating lung cancer, such as small cell lung cancer, non-small cell
CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
CC for diagnostic purposes and as targets for screening for therapeutic
CC compounds that modulate lung cancer, such as antibodies. Sequences
CC ABUS6408-ABUS6445 represent lung cancer-associated polypeptides of the
CC invention.
XX
SQ Sequence 243 AA;

```

```

Query Match 100.0%; Score 1303; DB 6; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MRPQGPASPORLGLLLLLLPASSASSEIRPKGKOKAQLRQREVDVLYNMCLOGPA 60
DB 1 MRPQGPASPORLGLLLLLLPASSASSEIRPKGKOKAQLRQREVDVLYNMCLOGPA 60
QY 61 GYPGRDGSFGANCIPGTGPIGKDGKGEKCECLRSFBSWTPNTYKQCSWSLNTYDIL 120
DB 61 GYPGRDGSFGANCIPGTGPIGKDGKGEKCECLRSFBSWTPNTYKQCSWSLNTYDIL 120

```

```

QY 121 GKIACTETKMSNSALVLFSGSLRLKCRNACCORMYFTFNAGCSGPLPEATIIYDQ 180
DB 121 GKIACTETKMSNSALVLFSGSLRLKCRNACCORMYFTFNAGCSGPLPEATIIYDQ 180
QY 181 GSPENNSTINIHRTSSVGLCEGIGAGLVDAIWIWGTCSDPKGDASTGWSVSRIIEE 240
DB 181 GSPENNSTINIHRTSSVGLCEGIGAGLVDAIWIWGTCSDPKGDASTGWSVSRIIEE 240
QY 241 LPK 243
DB 241 LPK 243

```

```

RESULT 9
ADB80510
ID ADB80510 standard; protein; 243 AA.
XX
AC ADB80510;
XX
DT 04-DEC-2003 (first entry)
XX
DE Ovarian cancer-associated protein #37.
XX
DE cytostatic; gene therapy; vaccine; ovarian cancer; diagnosis;
KW post-operative chemotherapy; radiation therapy; tumour prognosis;
KW pre-cancerous lesion detection.
XX
OS Homo sapiens.
XX
PN WO2002102235-A2.
XX
PD 27-DEC-2002.
XX
PF 18-JUN-2002; 2002WO-US019297.
XX
PR 18-JUN-2001; 2001US-0299234P.
XX
PR 27-AUG-2001; 2001US-0315287P.
XX
PR 05-SEP-2001; 2001US-0317544P.
XX
PR 13-NOV-2001; 2001US-0350666P.
XX
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
PI Mack DH, Gish KC;
XX
DR WPI; 2003-167431/16.
XX
DR N-PSDB; ADB80509.
XX
PT Detecting an ovarian cancer-associated transcript in a cell from a
PT patient, comprising contacting a biological sample from the patient with a
PT polynucleotide that hybridizes to an ovarian cancer gene.
XX
PS Claim 13; Page 299; 332pp; English.
XX

```

```

CC The invention relates to a method of detecting an ovarian cancer-
CC associated transcript in a cell from a patient, by contacting a
CC biological sample from the patient with a polynucleotide that selectively
CC hybridizes to a sequence at least 80% identical to any of one of 80
CC nucleic acid sequences given in the specification. The method is useful
CC in diagnosing ovarian cancer and in identifying and using agents and/or
CC targets that inhibit ovarian cancer. The nucleic acid molecule,
CC polypeptide and the antibody may also be used in detecting ovarian
CC cancer; monitoring and early detection of relapse following treatment,
CC monitoring response to therapy, selecting patients for post-operative
CC chemotherapy or radiation therapy, in selecting mode of therapy,
CC determining tumour prognosis, early detection of pre-cancerous lesions,
CC and as vaccines. This sequence corresponds to one of the proteins used
CC for the detection method of the invention.
XX
SQ Sequence 243 AA;

```

```

Query Match 100.0%; Score 1303; DB 7; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;

```

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAAPORLRLGLLLLLQLPAPSSASEIPKQKQAKLQREVVLDLYNMGCLQGPA 60  
DB 1 MRPGPAAPORLRLGLLLLLQLPAPSSASEIPKQKQAKLQREVVLDLYNMGCLQGPA 60  
QY 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKEGCELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
DB 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKEGCELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
QY 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQRYFTFNGAECGSPPIEAIITLDQ 180  
DB 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQRYFTFNGAECGSPPIEAIITLDQ 180  
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSRITIEE 240  
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSRITIEE 240  
QY 241 LPK 243  
DB 241 LPK 243

RESULT 10  
ADN39855  
ID ADN39855 standard; protein: 243 AA.

AC ADN39855;  
XX  
XX 17-JUN-2004 (first entry)  
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C225.

XX Human; differential expression; cancer; angiogenic disorder;  
XX fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
XX inflammatory disease; autoimmune disease;  
XX retinal neovascularisation syndrome; scarring; uterine fibroid;  
XX wound healing; diagnosis; prognosis; drug screening; drug targeting;  
XX vulnery; gene therapy; cytostatic; cardiant; immunomodulatory;  
XX  
XX Homo sapiens.  
XX OS  
XX PN MO2003042661-A2.  
XX  
XX 22-MAY-2003.  
XX PD  
XX PF 13-NOV-2002; 2002WO-US036810.  
XX  
XX 13-NOV-2001; 2001US-0350666P.  
XX PR 21-NOV-2001; 2001US-0332464P.  
XX PR 29-NOV-2001; 2001US-0334393P.  
XX PR 03-DEC-2001; 2001US-0335394P.  
XX PR 14-DEC-2001; 2001US-0340376P.  
XX PR 08-JAN-2002; 2002US-0347211P.  
XX PR 10-JAN-2002; 2002US-0347349P.  
XX PR 08-FEB-2002; 2002US-0355250P.  
XX PR 13-FEB-2002; 2002US-0356714P.  
XX PR 20-FEB-2002; 2002US-0359077P.  
XX PR 29-MAR-2002; 2002US-0368809P.  
XX PR 04-APR-2002; 2002US-0370110P.  
XX PR 12-APR-2002; 2002US-0372246P.  
XX PR 05-JUN-2002; 2002US-0386644P.  
XX PR 16-JUL-2002; 2002US-0396839P.  
XX PR 22-JUL-2002; 2002US-0397715P.  
XX PR 22-JUL-2002; 2002US-0397845P.  
XX PR 09-SEP-2002; 2002US-0409450P.  
XX  
XX (EOSB-) EOS BIOTECHNOLOGY INC.  
XX PA  
XX Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;  
XX Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;

DR WPI: 2003-468649/44.  
DR N-PSDB; ADN39855.  
PT Determining the presence or absence of a pathological cell in a patient,  
PT useful for diagnosing, prognosing or treating cancer, comprises detecting  
PT a nucleic acid in a biological sample.  
XX  
XX Claim 12, SEQ ID NO C225; 1385bp; English.  
XX  
XX The invention relates to nucleic acids and proteins (ADN39855-ADN40064)  
XX whose expression is upregulated or downregulated in specific cancers or  
XX other diseases such as angiogenic or fibrotic disorders, and to methods  
XX of determining the presence or absence of a pathological cell in a  
XX patient by detecting a nucleic acid at least 80% identical to those of  
XX the invention or by detecting a polypeptide of the invention. The  
XX invention also relates to expression vectors and host cells comprising a  
XX nucleic acid of the invention; antibodies which specifically bind a  
XX polypeptide of the invention; use of such antibodies for drug targeting;  
XX and methods of screening for modulators of activity or expression of the  
XX polypeptides and nucleic acids. The nucleic acids, polypeptides, and  
XX antibodies and methods are useful for diagnosing, prognosing and treating  
XX cancer and other conditions such as psoriasis, ischaemia, heart disease,  
XX atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
XX neovascularisation syndromes, scarring and uterine fibroids. They may  
XX also be useful in wound healing and in contraception. The present  
XX sequence represents a polypeptide of the invention.  
XX  
XX Sequence 243 AA;  
SQ

Query Match 100.0%; Score 1303; DB 7; Length 243;  
Best Local Similarity 100.0%; Pred. No. 9e-121;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAAPORLRLGLLLLLQLPAPSSASEIPKQKQAKLQREVVLDLYNMGCLQGPA 60  
DB 1 MRPGPAAPORLRLGLLLLLQLPAPSSASEIPKQKQAKLQREVVLDLYNMGCLQGPA 60  
QY 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKEGCELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
DB 61 GVPGRDSSPGANGIPGTPTGIPGRDGFKEGKEGCELRSEFSESWTPNPKQCSWSSLNVTGIDL 120  
QY 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQRYFTFNGAECGSPPIEAIITLDQ 180  
DB 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQRYFTFNGAECGSPPIEAIITLDQ 180  
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSRITIEE 240  
DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSRITIEE 240  
QY 241 LPK 243  
DB 241 LPK 243

RESULT 11  
ADN38732  
ID ADN38732 standard; protein: 243 AA.

AC ADN38732;  
XX  
XX 17-JUN-2004 (first entry)  
DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:50.

XX Human; differential expression; cancer; angiogenic disorder;  
XX fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;  
XX inflammatory disease; autoimmune disease;  
XX retinal neovascularisation syndrome; scarring; uterine fibroid;  
XX wound healing; diagnosis; prognosis; drug screening; drug targeting;  
XX vulnery; gene therapy; cytostatic; cardiant; immunomodulatory;  
XX  
XX Homo sapiens.

```

XX PN MO20030426661-A2.
XX PD
XX PF 22-MAY-2003.
XX PF 13-NOV-2002; 2002WO-US036810.
XX PR 13-NOV-2001; 2001US-0350666P.
XX PR 21-NOV-2001; 2001US-0332464P.
XX PR 29-NOV-2001; 2001US-0334393P.
XX PR 03-DEC-2001; 2001US-0335394P.
XX PR 14-DEC-2001; 2001US-0340376P.
XX PR 08-JAN-2002; 2002US-0347211P.
XX PR 10-JAN-2002; 2002US-0347349P.
XX PR 08-FEB-2002; 2002US-0355250P.
XX PR 13-FEB-2002; 2002US-0356714P.
XX PR 20-FEB-2002; 2002US-0359077P.
XX PR 29-MAR-2002; 2002US-036809P.
XX PR 04-APR-2002; 2002US-0370110P.
XX PR 12-APR-2002; 2002US-0372246P.
XX PR 05-JUN-2002; 2002US-038614P.
XX PR 16-JUL-2002; 2002US-0396839P.
XX PR 22-JUL-2002; 2002US-039775P.
XX PR 22-JUL-2002; 2002US-0397845P.
XX PR 09-SEP-2002; 2002US-0409450P.
XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX PI Afar D, Aziz N, Gineburg WM, Gish KC, Glynn R, Hevezi PA;
XX PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnick A;
XX DR N-PSDB; ADN38731.
XX DR WPI; 2003-468649/44.
XX PT Determining the presence or absence of a pathological cell in a patient,
XX PT useful for diagnosing, prognosing or treating cancer, comprises detecting
XX PT a nucleic acid in a biological sample.
XX PS Claim 12; SEQ ID NO 50; 1385pp; English.
XX CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
XX CC whose expression is upregulated or downregulated in specific cancers or
XX CC other diseases such as angiogenic or fibrotic disorders, and to methods
XX CC of determining the presence or absence of a pathological cell in a
XX CC patient by detecting a nucleic acid at least 80% identical to those of
XX CC the invention or by detecting a polypeptide of the invention. The
XX CC invention also relates to expression vectors and host cells comprising a
XX CC nucleic acid of the invention; antibodies which specifically bind a
XX CC polypeptide of the invention; use of such antibodies for drug targeting;
XX CC and methods of screening for modulators of activity or expression of the
XX CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
XX CC antibodies and methods are useful for diagnosing, prognosing and treating
XX CC cancer and other conditions such as psoriasis, ischemia, heart disease,
XX CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
XX CC neovascularization syndromes, scarring and uterine fibroids. They may
XX CC also be useful in wound healing and in contraception. The present
XX CC sequence represents a polypeptide of the invention.
XX SQ Sequence 243 AA;
XX
Query Match 100.0%; Score 1303; DB 7; Length 243;
Best Local Similarity 100.0%; Pred. No. 9e-121;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRQGPASAPQRLGILLILLLOLPAPSSASEIPKQKQQLRQREVVLDLYNMCLOQPA 60
DB 1 MRQGPASAPQRLGILLILLLOLPAPSSASEIPKQKQQLRQREVVLDLYNMCLOQPA 60
QY 61 GVGGRDGPAGANGIPGPGIPGRDGRKGEKGECLARSFESWPNPKQCSWSSLYNGIDL 120
DB 61 GVGGRDGPAGANGIPGPGIPGRDGRKGEKGECLARSFESWPNPKQCSWSSLYNGIDL 120
QY 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYTFNGAGCSGPLPDAIYLDQ 180

```

```

DB 121 GKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYTFNGAGCSGPLPDAIYLDQ 180
QY 181 GSPENNSTINHRTSSVBGLCEGAGLVDAIIVGTCSDPKPGDASTGMSVSRILIEE 240
DB 181 GSPENNSTINHRTSSVBGLCEGAGLVDAIIVGTCSDPKPGDASTGMSVSRILIEE 240
QY 241 LPK 243
DB 241 LPK 243
RESULT 12
ADL70254
ID ADL70254 standard; protein; 243 AA.
XX ADL70254;
AC ADL70254;
XX 20-MAY-2004 (first entry)
DT
XX LBFJ301 polypeptide, associated with stomach cancer.
DE LBFJ301; human; stomach cancer; diagnosis; cytostatic.
XX LBFJ301; human; stomach cancer; diagnosis; cytostatic.
XX Homo sapiens.
XX MO2004016636-A1.
XX 26-FEB-2004.
XX 14-AUG-2003; 2003WO-KR001653.
XX 14-AUG-2002; 2002US-0402904P.
XX 20-AUG-2002; 2002US-0404408P.
XX 23-AUG-2002; 2002US-0405304P.
XX 28-OCT-2002; 2002US-0421582P.
XX (GLDS ) LG LIFE SCI LTD.
XX Koh SS, Liu Q, Chung H, Zeng W, Lee B, Song SY;
XX WPI; 2004-203757/19.
XX N-PSDB; ADL70253.
XX PT New nucleic acid molecule associated with stomach cancer, useful for
XX PT preparing a composition for diagnosing or treating cancers such as
XX PT stomach cancer.
XX Claim 24; SEQ ID NO 2; 146pp; English.
XX CC The present sequence is that of the polypeptide encoded by cDNA clone
XX CC AD12. This clone corresponds to an mRNA species that is differentially
XX CC expressed in cancerous stomach tissue compared to healthy stomach tissue.
XX CC Clones AD12 and CH4 ADL70255 are splice variants of a gene designated
XX CC LBFJ301. Expression of this gene is significantly up-regulated (13.75-
XX CC fold) in gastric carcinoma samples compared to samples from healthy
XX CC stomach tissue. Up-regulation of LBFJ301 may therefore be diagnostic for
XX CC stomach cancer. The predicted proteins encoded by AD12 and CH4 ADL70256
XX CC are identical for the first 124 amino acids, while the last 13 amino
XX CC acids of the CH4 protein are unique. Termination of the CH4 protein is
XX CC produced by a 45 bp insertion which introduces a stop codon in the open
XX CC reading frame. The proteins are weakly similar to the chymotrypsin serine
XX CC protease family signature and the NUDIX hydrolase family signature. The
XX CC nucleic acids and proteins of the invention can be used as diagnostic
XX CC agents or markers to detect stomach cancer or to monitor the progression
XX CC of stomach cancer in a sample. The proteins can be used to identify
XX CC agents which modulate protein level or activity, to identify binding
XX CC partners, to raise antibodies, as therapeutic targets, and as diagnostic
XX CC agents or markers of stomach cancer and other hyperplastic diseases. A
XX CC claimed method for diagnosing a disease state in a subject comprises
XX CC determining the level of expression of the nucleic acid, where the
XX CC disease state is stomach cancer, advanced gastric cancer or a malignant
XX CC neoplasm occurring in soft tissue, bone, breast, cervix, colon,

```





XX Homo sapiens.  
 OS  
 XX  
 XX WO200153455-A2.  
 PN  
 XX  
 PD 26-JUL-2001.  
 XX  
 XX  
 PF 22-DEC-2000; 2000MO-US035017.  
 XX  
 XX 23-DEC-1999; 99US-00471275.  
 PR 21-JAN-2000; 2000US-00488725.  
 PR 25-APR-2000; 2000US-00552317.  
 XX  
 XX (HYSEQ) HYSEQ INC.  
 PA  
 XX Tang YT, Liu C, Drmanac RT;  
 PI  
 XX WPI; 2001-457603/49.  
 DR  
 XX N-PSDB; AAH99687.  
 XX  
 PT Isolated human polynucleotides encoding polypeptides, useful for the  
 PT treatment and diagnosis of e.g. cancer, ulcers and HIV infection.  
 XX  
 XX Claim 20; Page 260; 1217pp; English.  
 XX  
 CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
 CC AAM25963. The proteins can have activities based on the tissues and cells  
 CC they are expressed in, such as: antiinflammatory; antirheumatic;  
 CC antiarthritic; immunosuppressive; antibacterial; endocrine; cardiac;  
 CC central nervous system; virolytic; anti-HIV; fungicide; antimutagen;  
 CC cardiovascular; antihaemic; antiaggregant; haemostatic; vulnerrary;  
 CC antitumor; osteopathic; dermatological; antiallergic; antiaesthetic;  
 CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
 CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
 CC encoding them can be used in gene therapy, antisense therapy and vaccine  
 CC production. The proteins and polynucleotides are useful for screening for  
 CC agonists or antagonists of a protein and for the treatment and diagnosis  
 CC of disorders associated with the activity of a protein e.g. inflammation,  
 CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
 CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
 CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
 CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, allergic  
 CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
 CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
 CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
 CC neurological disorders  
 CC  
 XX  
 SQ Sequence 278 AA;  
 Query Match 100.0%; Score 1303; DB 4; Length 278;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-120;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MRQGPAAAPQRLRGLLLLLQLPAPSSASERPKGKQALRQREVDVLYNMCLOGPA 60  
 DB 36 MRQGPAAAPQRLRGLLLLLQLPAPSSASERPKGKQALRQREVDVLYNMCLOGPA 95  
 QY 61 GVGGRGSSPGANGIPGTPIGSDGKGEKGECLRSFESSWTPNTYKQCSWSSLYATGIDL 120  
 DB 96 GVGGRGSSPGANGIPGTPIGSDGKGEKGECLRSFESSWTPNTYKQCSWSSLYATGIDL 155  
 QY 121 GKLAECFTKMRNSALRVLFSGSLRLKCNACCOMRYFPFNABCSGPIPIAIIYLDQ 180  
 DB 156 GKLAECFTKMRNSALRVLFSGSLRLKCNACCOMRYFPFNABCSGPIPIAIIYLDQ 215  
 QY 181 GSPFNASTINIRHTSSVEGLCEGIGAGLVDAIVWGTCSDPYKGDASTGMNSYSRIIIE 240  
 DB 216 GSPFNASTINIRHTSSVEGLCEGIGAGLVDAIVWGTCSDPYKGDASTGMNSYSRIIIE 275  
 QY 241 LPK 243  
 DB 276 LPK 278

RESULT 15  
 ABG96338  
 ID ABG96338 standard; protein; 278 AA.  
 XX  
 XX  
 AC ABG96338;  
 XX  
 XX  
 DT 11-DEC-2002 (first entry)  
 XX  
 XX  
 DE Human ovarian cancer marker M138.  
 XX  
 XX Human ovarian cancer; marker; cancer; familial history; brain disorder;  
 KW central nervous system disorder; bacterial meningitis; viral meningitis;  
 KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;  
 KW brain herniation; inflammation; encephalitis; testicular disorder;  
 KW nonbuberculous granulomatous orchitis; connective tissue disorder;  
 KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;  
 KW histological type; carcinogenic; ovarian cancer marker.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200271928-A2.  
 PN  
 XX  
 PD 19-SEP-2002.  
 XX  
 XX  
 PF 14-MAR-2002; 2002MO-US007826.  
 XX  
 XX 14-MAR-2001; 2001US-0276025P.  
 PR 14-MAR-2001; 2001US-0276026P.  
 PR 10-AUG-2001; 2001US-0311732P.  
 PR 19-SEP-2001; 2001US-0323580P.  
 PR 26-SEP-2001; 2001US-0324967P.  
 PR 26-SEP-2001; 2001US-0325102P.  
 PR 26-SEP-2001; 2001US-0325149P.  
 XX  
 XX (MILL-) MILLENNIUM PHARM INC.  
 PA  
 XX Monahan JE, Gamavarrapu M, Hoersch S, Kamatkar S, Kovatis SG;  
 PI Meyers RE, Morrissey MP, Olandt PJ, Sen A, Vieby PO, Mills GB;  
 PI Bast RC, Lu K, Schmandt RB, Zhao X, Glatt K;  
 XX  
 XX WPI; 2002-723277/78.  
 DR N-PSDB; ABS76434.  
 XX  
 PT Assessing whether a patient is afflicted with ovarian cancer, useful in  
 PT assessing the stage or progression of the disease, comprises comparing  
 PT the expression level of a cancer marker in a sample from a patient and  
 XX from a non cancer patient.  
 XX  
 PS Disclosure; Page 249-250; 481pp; English.  
 XX  
 XX The present invention relates to a new method for assessing whether a  
 CC patient is afflicted with ovarian cancer. The method involves comparing  
 CC the expression level of a marker in a patient sample and the normal level  
 CC of expression of the marker in a control non-ovarian cancer sample, where  
 CC the marker is selected from 363 cancer markers described in the  
 CC specification. The method of the invention is useful in diagnosing or  
 CC characterising cancer, in detecting the presence of cancer as early as  
 CC possible, and the recurrence of ovarian cancer. The method may also be of  
 CC particular use with patients having an enhanced risk of developing  
 CC ovarian cancer (e.g. patients having a familial history of ovarian  
 CC cancer). The cancer markers may be used in the management and treatment  
 CC of e.g. brain and central nervous system disorders (e.g. bacterial and  
 CC viral meningitis, Alzheimer's disease or Parkinson's disease), brain  
 CC disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),  
 CC inflammations (e.g. bacterial or viral meningitis or encephalitis),  
 CC testicular disorders (e.g. nonbuberculous granulomatous orchitis),  
 CC connective tissue disorders, or heart disorders (e.g. ischaemic heart  
 CC disease or atherosclerosis). The compositions and methods may also be  
 CC used in assessing the histological type of neoplasm associated with  
 CC ovarian cancer, monitoring the progression of ovarian cancer, determining  
 CC whether ovarian cancer has metastasized or is likely to metastasize,  
 CC selecting a composition for inhibiting ovarian cancer, assessing the





This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:56 ; Search time 24.4588 Seconds  
(without alignments)  
741.643 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303  
1 MRPGPAPSPQRRLGILLLL.....GDASTGMSVSRRIIEELRK 243

Sequence: BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	4	US-09-692-081-4
2	1300	99.8	243	4	US-09-489-847-205
3	1294	99.3	243	4	US-09-834-759-514
4	1294	99.3	278	4	US-09-834-759-515
5	1225	94.0	245	4	US-09-692-081-2
6	1225	94.0	277	4	US-09-692-081-5
7	1109.5	85.1	276	4	US-09-205-258-958
8	1046	80.3	197	4	US-09-834-759-516
9	1046	80.3	232	4	US-09-834-759-517
10	356	27.3	66	4	US-09-205-258-962
11	286	22.7	52	4	US-09-205-258-961
12	270	20.7	51	4	US-09-205-258-963
13	248	19.0	46	4	US-09-834-759-518
14	220.5	16.9	52	4	US-09-205-258-960
15	196	15.0	93	4	US-09-489-847-240
16	196	15.0	93	4	US-09-489-847-362
17	141	10.8	26	4	US-09-834-759-519
18	131.5	10.1	246	2	US-08-463-911-4
19	127.5	9.8	1057	3	US-08-931-820-1
20	127.5	9.8	1461	4	US-09-585-887-9
21	127.5	9.8	1461	4	US-09-289-578-9
22	127.5	9.8	1464	4	US-09-331-347C-21
23	125	9.6	313	4	US-09-949-016-9265
24	125	9.6	1670	4	US-09-949-016-5883
25	124.5	9.6	128	3	US-09-227-357-190
26	124.5	9.6	1516	4	US-09-949-016-8209
27	124	9.5	623	3	US-09-029-348-3

28	124	9.5	626	3	US-09-029-348-2	Sequence 2, Appl
29	122.5	9.4	1268	4	US-09-949-016-7487	Sequence 7487, Ap
30	121.5	9.3	492	4	US-08-468-996-12	Sequence 12, Appl
31	121	9.3	96	4	US-09-513-999C-4206	Sequence 4206, Ap
32	121	9.3	208	4	US-09-471-276-823	Sequence 823, App
33	121	9.3	245	4	US-09-800-729-151	Sequence 151, App
34	120	9.2	245	4	US-09-552-225A-4	Sequence 4, Appl
35	120	9.2	245	4	US-09-552-204A-4	Sequence 4, Appl
36	119.5	9.2	310	3	US-09-219-849-47	Sequence 47, Appl
37	119.5	9.2	595	3	US-09-219-849-48	Sequence 48, Appl
38	119.5	9.2	595	3	US-09-219-849-50	Sequence 50, Appl
39	119.5	9.2	684	3	US-08-555-669-12	Sequence 12, Appl
40	119.5	9.2	684	3	US-09-073-663-12	Sequence 12, Appl
41	119.5	9.2	755	4	US-09-919-497-57	Sequence 57, Appl
42	119.5	9.2	822	3	US-09-219-849-49	Sequence 49, Appl
43	119	9.1	258	4	US-09-976-594-815	Sequence 815, Appl
44	119	9.1	1806	4	US-09-919-497-56	Sequence 56, Appl
45	118.5	9.1	1694	1	US-08-494-168-2	Sequence 2, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-692-081-4
; Sequence 4, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-692-081-4

Query Match      100.0%; Score 1303; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 6.5e-125;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPGPAPSPQRRLGILLLLQLPAPSSASEIPKQKQAKLROREVDLYNGMCIQGA 60
Db      1 MRPGPAPSPQRRLGILLLLQLPAPSSASEIPKQKQAKLROREVDLYNGMCIQGA 60

QY      61 GVPGRDGFANGIPGTGIPGRDGFKGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120
Db      61 GVPGRDGFANGIPGTGIPGRDGFKGKGECLRESFEESWTPNYKQCSWSLNYGIDL 120

QY      121 GIIARCTPTKMSNSALRYLFGSLRLKCRNACQQRWTFMGAGSGGLPIEATLYLDO 180
Db      121 GIIARCTPTKMSNSALRYLFGSLRLKCRNACQQRWTFMGAGSGGLPIEATLYLDO 180

QY      181 GSPENNSTINIRTSVEGLCEGAGLVDAVIMWGTCSDYDKGDASTGMSVSRRIIRE 240
Db      181 GSPENNSTINIRTSVEGLCEGAGLVDAVIMWGTCSDYDKGDASTGMSVSRRIIRE 240

QY      241 LPK 243
Db      241 LPK 243

RESULT 2
US-09-489-847-205
; Sequence 205, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
```

```

? TITLE OF INVENTION: 98 Human Secreted Proteins
? FILE REFERENCE: P2031p1
? CURRENT APPLICATION NUMBER: US/09/489, 847
? CURRENT FILING DATE: 2000-01-24
? EARLIER APPLICATION NUMBER: PCT/US99/17130
? EARLIER FILING DATE: 1999-07-29
? EARLIER APPLICATION NUMBER: 60/094, 657
? EARLIER FILING DATE: 1998-07-30
? EARLIER APPLICATION NUMBER: 60/095, 486
? EARLIER FILING DATE: 1998-08-05
? EARLIER APPLICATION NUMBER: 60/096, 319
? EARLIER FILING DATE: 1998-08-12
? EARLIER APPLICATION NUMBER: 60/095, 454
? EARLIER FILING DATE: 1998-08-06
? EARLIER APPLICATION NUMBER: 60/095, 455
? EARLIER FILING DATE: 1998-08-06
? NUMBER OF SEQ ID NOS: 376
? SOFTWARE: PatentIn Ver. 2.0
? SEQ ID NO 205
? LENGTH: 243
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-09-489-847-205

```

Query Match	99.8%	Score 1300	DB 4	Length 243
Best Local Similarity	99.6%	Pred. No. 1.3e-124		
Matches 242	Conservative	1	Mismatches 0	Indels 0
				Gaps 0

Qy	1	MRQGPAAASQRLRGILLILLILLQLPAPSSASIPKQKQALRREVVLDYNGMCLOGPA	60
Db	1	MRQGPAAASQRLRGILLILLILLQLPAPSSASIPKQKQALRREVVLDYNGMCLOGPA	60
Qy	61	GVPRGDSPGANGIPETPGIPGRDGFKEGKEGCLNESTFEESWTNYKQSSWNLNYGIDL	120
Db	61	GVPRGDSPGANGIPETPGIPGRDGFKEGKEGCLNESTFEESWTNYKQSSWNLNYGIDL	120
Qy	121	GKLAECTFTMRNSALRYLFGSGLRLKCRNACCCORWFTFNGAGCSGPLPTEAIIYLDQ	180
Db	121	GKLAECTFTMRNSALRYLFGSGLRLKCRNACCCORWFTFNGAGCSGPLPTEAIIYLDQ	180
Qy	181	GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKPDASTGWNVSRIIIEE	240
Db	181	GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWTGTCSDYPKPDASTGWNVSRIIIEE	240
Qy	241	LPK 243	
Db	241	LPK 243	

RESULT 3  
 US-09-834-759-514  
 Sequence 514, Application US/09834759  
 Patent No. 6680197  
 GENERAL INFORMATION:  
 APPLICANT: Jiang, Yugu  
 APPLICANT: Dillon, David C.  
 APPLICANT: Mitcham, Jennifer L.  
 APPLICANT: Xu, Jiaqichun  
 APPLICANT: Harlocker, Susan L.  
 APPLICANT: Hepler, William T.  
 APPLICANT: Henderson, Robert A.  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 FILE REFERENCE: 210121.470C9  
 CURRENT APPLICATION NUMBER: US/09/834,759  
 CURRENT FILING DATE: 2001-04-13  
 NUMBER OF SEQ ID NOS: 547  
 SOFTWARE: FastSeq for Windows Version 3.0  
 SEQ ID NO 514  
 LENGTH: 243  
 TYPE: PRN  
 ORGANISM: Homo sapiens  
 US-09-834-759-514

Query Match	99.3%	Score 1294;	DB 4;	Length 243;
Best Local Similarity	99.6%	Pred. NO. 5.4e-124;		
Matches 242; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

Qy	2	MRQGGAAAPQRLRGILLILLILLQLPAPSSASEIFKQKQKQRLRREVVLDLYNMGCLQSPA	60
Db	1	MRQGGAAAPQRLRGILLILLILLQLPAPSSASEIFKQKQKQRLRREVVLDLYNMGCLQSPA	60
Qy	61	GVPGRDGSFGANGIFCTPGIPIGRDGFKEKECECLRESFEBSWTNYYQCSWSSLNTGIDL	120
Db	61	GVPGRDGSFGAAVVIPTGPPIPIGRDGFKEKECECLRESFEBSWTNYYQCSWSSLNTGIDL	120
Qy	121	GKIAECTFPKMSNSALRYLFSGSRLTKCRNACCOMRYFTFNGAEGSGPPIEAILYLDQ	180
Db	121	GKIAECTFPKMSNSALRYLFSGSRLTKCRNACCOMRYFTFNGAEGSGPPIEAILYLDQ	180
Qy	181	GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDYPKDASTGMNSVSRIIIEE	240
Db	181	GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDYPKDASTGMNSVSRIIIEE	240
Qy	241	LPK	243
Db	241	LPK	243

```

RESULT 4
US-09-834-759-515
; Sequence 515, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515

```

Query Match	99.3%	Score 1294	DB 4	Length 278
Best Local Similarity	99.6%	Pred. No. 6.5e-124		
Matches 242	Conservative 0	Mismatches 1	Indels 0	Gaps 0
Qy	1	MRPGAPASPORLGLLILLILLQLPASSASAEIPKQKQALRQREVVDLYNMGCLQGPA	60	
Db	36	MRPGAPASPORLGLLILLILLQLPAPSSASEIPKQKQALRQREVVDLYNMGCLQGPA	95	
Qy	61	GVPRDQSGPGANGIPGTPIPGRDGFKEKECELCRESFEESWTTPNYKCCSSSLNYGIDL	120	
Db	96	GVPRDQSGPGANVIPGTPIPGRDGFKEKECELCRESFEESWTTPNYKCCSSSLNYGIDL	155	
Qy	121	GKIAECTTKRNSALRVLPFSGSLRLCKRNACCORWFETENGACSGPLPLEATIIYDQ	180	
Db	156	GKIAECTTKRNSALRVLPFSGSLRLCKRNACCORWFTENGACSGPLPLEATIIYDQ	215	
Qy	181	GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIIIEE	240	
Db	216	GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGMSVSRIIIEE	275	
Qy	241	LPK 243		



```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 958
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-958
```

```

Query Match      85.1%; Score 1109.5; DB 4; Length 276;
Best Local Similarity 97.1%; Pred. No. 4.4e-105;
Matches 204; Conservative 3; Mismatches 2; Indels 1; Gaps 1;
```

```

QY 34 PKGKQAKLRORREVDLYNMCLOQPGVGRDGPANGIPCTPGIPGRDGFKEKGRG 93
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 68 PRG-SRRIRKREVDLYNMCLOQPGVGRDGPANGIPCTPGIPGRDGFKEKGRG 126
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
QY 94 LRSPFESWTPTYKQCSWSLNYGIDLGKIAECTFTKMSNSALRYLFGSGLRLKCRNAC 153
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 127 LRSPFESWTPTYKQCSWSLNYGIDLGKIAECTFTKMSNSALRYLFGSGLRLKCRNAC 186
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
QY 154 CORWYFTFNGAECGSGPLPIEAIYYLDQSGPEMNSTINIHRTSSVEGLCEGIGALVDVAI 213
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 187 CORWYFTFNGAECGSGPLPIEAIYYLDQSGPEMNSTINIHRTSSVEGLCEGIGALVDVAI 246
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
QY 214 WVGTCSDYPRKGDASTGMNSVSRITIEELPK 243
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 247 WVGTCSDYPRKGDASTGMNSVSRITIEELPK 276
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
```

```

RESULT 8
US-09-834-759-516
; Sequence 516, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
```

```

; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaqun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Heppler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 516
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-516
```

```

Query Match      80.3%; Score 1046; DB 4; Length 197;
Best Local Similarity 99.5%; Pred. No. 8.5e-99;
Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 MRPGPAPSPQRLRGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNMCLOGPA 60
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 1 MRPGPAPSPQRLRGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNMCLOGPA 60
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
QY 61 GVPGRDGPANGIPCTPGIPGRDGFKEKGRGCLRESFEESWTPTYKQCSWSLNYGIDL 120
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 61 GVPGRDGPANGIPCTPGIPGRDGFKEKGRGCLRESFEESWTPTYKQCSWSLNYGIDL 120
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
QY 121 GKIACTFTKMSNSALRYLFGSGLRLKCRNACCORWYFTFNGAECGSLPIEAIYYLDQ 180
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 121 GKIACTFTKMSNSALRYLFGSGLRLKCRNACCORWYFTFNGAECGSLPIEAIYYLDQ 180
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
QY 181 GSPENSTINIHRTSSV 197
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 181 GSPENSTINIHRTSSV 197
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
```

```

RESULT 9
US-09-834-759-517
; Sequence 517, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaqun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Heppler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 517
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-517
```

```

Query Match      80.3%; Score 1046; DB 4; Length 232;
Best Local Similarity 99.5%; Pred. No. 1.1e-98;
Matches 196; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```

QY 1 MRPGPAPSPQRLRGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNMCLOGPA 60
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
DB 36 MRPGPAPSPQRLRGILLILLQLPAPSSASEIPKQKQAKLRORREVDLYNMCLOGPA 95
   |.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|.:|
```

```
QY      61 GVPGRGSGPGANIPGTGPGDGRGKGEKCELRSPFSWTPNTKQCSWSLNYGIDL 120
DB      96 GVPGRGSGSPANIPGTGPGDGRGKGEKCELRSPFSWTPNTKQCSWSLNYGIDL 155
QY     121 GKIAECTFTMRNSALRVLFSGSLRLKCRNACCORWYTFNGAEGSGPLPIEAIYYLDQ 180
DB     156 GKIAECTFTMRNSALRVLFSGSLRLKCRNACCORWYTFNGAEGSGPLPIEAIYYLDQ 215
QY      181 GSPEMNSTINIRHTSSV 197
DB     216 GSPEMNSTINIRHTSSV 232

RESULT 10.
US-09-205-258-962
; Sequence 962, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 962
; LENGTH: 66
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-962

Query Match      27.3%; Score 356; DB 4; Length 66;
Best Local Similarity 100.0%; Pred. No. 5.5e-29;
Matches 66; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      128 FTRKMSNSALRVLFSGSLRLKCRNACCORWYTFNGAEGSGPLPIEAIYYLDQSPENMS 187
DB      1  FTRKMSNSALRVLFSGSLRLKCRNACCORWYTFNGAEGSGPLPIEAIYYLDQSPENMS 60
QY      188 TINHR 193
DB      61 TINHR 66

RESULT 11
US-09-205-258-961
; Sequence 961, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
```

```
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,875
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,374
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,917
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,949
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,974
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,883
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,897
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,898
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,962
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,963
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,877
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,878
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/070,923
EARLIER FILING DATE: 1997-12-18
EARLIER APPLICATION NUMBER: 60/092,921
EARLIER FILING DATE: 1998-07-15
EARLIER APPLICATION NUMBER: 60/094,657
EARLIER FILING DATE: 1998-07-30
NUMBER OF SEQ ID NOS: 1227
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 961
LENGTH: 52
```

```
TYPE: PRT
ORGANISM: Homo sapiens
US-09-205-258-961

Query Match      22.7% Score 296; DB 4; Length 52;
Best Local Similarity 100.0%; Pred. No. 5.3e-23;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      77 TPGIRGRGFKKGEKCEKRESFESEWTPNYQCSWSLNYGIDKIAECTP 128
DB      1 TPGIRGRGFKKGEKCEKRESFESEWTPNYQCSWSLNYGIDKIAECTP 52

RESULT 12
US-09-205-258-963
Sequence 963, Application US/09205258
Patent No. 6525174
GENERAL INFORMATION:
APPLICANT: Young et al.
TITLE OF INVENTION: 207 Human Secreted Proteins
FILE REFERENCE: P200791
CURRENT APPLICATION NUMBER: US/09/205,258
CURRENT FILING DATE: 1998-12-04
EARLIER APPLICATION NUMBER: PCT/US98/11422
EARLIER FILING DATE: 1998-06-04
EARLIER APPLICATION NUMBER: 60/048,885
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,375
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,881
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,880
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,896
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,020
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,876
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,895
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,884
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,894
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,971
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,964
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,882
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,899
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,893
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,900
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,901
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,892
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,915
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,019
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,970
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,972
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/048,916
EARLIER FILING DATE: 1997-06-06
EARLIER APPLICATION NUMBER: 60/049,373
EARLIER FILING DATE: 1997-06-06
```



EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 963  
LENGTH: 51  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-205-258-963

Query Match 20.7%; Score 270; DB 4; Length 51;  
Best Local Similarity 100.0%; Pred. No. 2.3e-20;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 193 RTSSVEGLCEGIGALVDVAIWGTCSDPKGDASTGMSVSRRIIEELPK 243  
Db 1 RTSSVEGLCEGIGALVDVAIWGTCSDPKGDASTGMSVSRRIIEELPK 51

RESULT 13  
US-09-834-759-518  
Sequence 518, Application US/09834759  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C9  
CURRENT APPLICATION NUMBER: US/09/834,759  
FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 547  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 518  
LENGTH: 46  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-759-518

Query Match 19.0%; Score 248; DB 4; Length 46;

Best Local Similarity 100.0%; Pred. No. 3.6e-18;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 198 EGLCEGIGALVDVAIWGTCSDPKGDASTGMSVSRRIIEELPK 243  
Db 1 EGLCEGIGALVDVAIWGTCSDPKGDASTGMSVSRRIIEELPK 46

RESULT 14  
US-09-205-258-960  
Sequence 960, Application US/09205258  
Patent No. 6525174  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P1  
CURRENT APPLICATION NUMBER: US/09/205,258  
FILING DATE: 1998-12-04  
EARLIER APPLICATION NUMBER: PCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 960
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-960

```

```

Query Match      16.9%; Score 220.5; DB 4; Length 52;
Best Local Similarity 87.2%; Pred. No. 2.7e-15;
Matches 41; Conservative 3; Mismatches 2; Indels 1; Gaps 1;

```

```

QY      34 PRGKAKQLRQREVDLYNGMCLQSPAGVGRDPSGANGIPGTPI 80
DB      7 PRG-SKRLRQREVDLYNGMCLQSPAGVGRDPSGANGIPGTPI 52

```

```

RESULT 15
US-09-489-847-240
; Sequence 240, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031PI
; CURRENT FILING DATE: US/09/489,847
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 240
; LENGTH: 93
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-240

```

```

Query Match      15.0%; Score 196; DB 4; Length 93;

```

```

Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 MRPGSPASPORTLRLGLLLLLLLQLPASPSSASEIPKQKXA 40
DB      1 MRPGSPASPORTLRLGLLLLLLLQLPASPSSASEIPKQKXA 40

```

```

Search completed: August 4, 2005, 04:23:53
Job time : 25.4588 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:16 ; Search time 19.3765 Seconds

(without alignments)  
1206.653 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303  
Sequence: 1 MRQGPAPASQRLRGILL.....GDASTGMSVSRRIIEELK 243Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	131.5	10.1	246	2 S29328	complement subcomp
2	129	9.9	1752	2 A45407	collagen alpha 3 (I
3	128	9.8	289	2 T20177	hypothetical prote
4	127.5	9.8	1464	1 CGHUI5	collagen alpha 1 (I
5	126	9.7	327	2 T29031	hypothetical prote
6	125	9.7	754	2 A55267	collagen alpha 5 (I
7	125	9.6	888	2 S28791	collagen alpha 1 (X
8	125	9.6	1670	1 CGHUI3	collagen alpha 3 (I
9	124	9.5	1466	1 CGHUI7L	collagen alpha 1 (I
10	124	9.5	1763	2 S16366	collagen alpha 2 (I
11	123.5	9.5	311	2 T15268	hypothetical prote
12	123	9.4	304	2 T26185	hypothetical prote
13	123	9.4	304	2 T26184	hypothetical prote
14	123	9.4	1691	1 S220177	collagen alpha 5 (I
15	122.5	9.4	300	2 T24482	hypothetical prote
16	122.5	9.4	920	2 A45748	collagen alpha 1 (V
17	122.5	9.4	1549	2 T48103	type VII collagen
18	122	9.4	2944	2 A54849	collagen alpha 1 (V
19	121	9.3	178	2 A39762	collagen alpha 1 (X
20	121	9.3	325	2 T18594	hypothetical prote
21	120.5	9.2	428	2 T24769	hypothetical prote
22	120	9.2	245	1 C1HUC	complement subcomp
23	119.5	9.2	358	2 T26281	hypothetical prote
24	119.5	9.2	298	2 T27644	hypothetical prote
25	119.5	9.2	310	2 T29731	hypothetical prote
26	119.5	9.2	438	2 T31631	hypothetical prote
27	119.5	9.2	671	1 CGRTIS	collagen alpha 1 (I
28	119.5	9.2	1453	2 S21626	collagen alpha 1 (I
29	119.5	9.2	1669	1 CGHUI4B	collagen alpha 1 (I

30	119	9.1	298	2 JCI448	collagen col-34 -
31	119	9.1	299	2 T29956	hypothetical prote
32	119	9.1	1042	1 CGCHIS	collagen alpha 1 (I
33	119	9.1	1049	1 CGBO7S	collagen alpha 1 (I
34	119	9.1	1806	1 CGHUIE	collagen alpha 1 (X
35	118.5	9.1	177	2 S37749	collagen alpha 2 (X
36	118.5	9.1	290	2 T24586	hypothetical prote
37	118.5	9.1	330	2 S46657	collagen alpha 1 (X
38	118.5	9.1	488	2 A27353	collagen alpha 1 (I
39	118.5	9.1	1691	1 CGHUI6B	collagen alpha 6 (I
40	118	9.1	1669	1 CGMS4B	collagen alpha 1 (I
41	118	9.1	1744	2 S40091	collagen alpha 1 (I
42	117.5	9.0	323	2 A61396	collagen alpha 1 (I
43	117.5	9.0	1486	1 B40333	collagen alpha 1 (I
44	117.5	9.0	1492	2 A40333	collagen alpha 1 (I
45	117	9.0	248	1 LNHUP1	pulmonary surfacta

## ALIGNMENTS

## RESULT 1

S29328  
complement subcomponent C1q chain C - mouse

C/Species: Mus musculus (house mouse)

C/Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004

R/Accession: S29328  
R/Petry, F.; Reid, K.B.M.; Loos, M.

Eur. J. Biochem. 209, 129-134, 1992

A/Title: Isolation, sequence analysis and characterization of cDNA clones coding for th

cerebellin.

A/Reference number: S29328; PMID:93011118; PMID:1396691

A/Accession: S29328

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-246 &lt;PET&gt;

A/Cross-references: UNIPROT:002105; EMBL:X66295; NID:950228; PTDN:CAA6993.1; PTD:95022

C/Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom

F;122-245/Domain: complement C1q carboxyl-terminal homology &lt;C1Q&gt;

Query Match 10.1%; Score 131.5; DB 2; Length 246;

Best Local Similarity 38.5%; Pred. No. 0.00026;

Matches 35; Conservative 5; Mismatches 32; Indels 19; Gaps 2;

OY 5 GPASFORLRLGLLLLLLPAPSSASERPKQKQALRQREVVDLYNGKCLQGPAGVPG 64

DB 4 GPSQPCQCGICLLILFLALPLRSQAS-----ACGYGIPMGKMGAGAG 47

OY 65 RDGS---PGANGIPGPRGIPGRDGFGEKGE 92

DB 48 KDGHGGLGPKBPPIPAVPGTGPQKQGE 78

## RESULT 2

A45407  
collagen alpha 3 (IV) chain - sea urchin (Strongylocentrotus purpuratus)

C/Species: Strongylocentrotus purpuratus (purple urchin)

C/Date: 22-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004

R/Accession: A45407; A43903; A23940

J. Biol. Chem. 268, 5249-5254, 1993

A/Title: Complete primary structure of a sea urchin type IV collagen alpha chain and an

A/Reference number: A45407; PMID:93186842; PMID:844699

A/Accession: A45407

A/Status: preliminary; not compared with conceptual translation

A/Molecule type: nucleic acid

A/Residues: 1-1752 &lt;EXP&gt;

A/Cross-references: UNIPROT:Q26312

A/Note: Sequence extracted from NCBI backbone (NCBIP:126841)

R/Wessel, G.M.; Etkin, M.; Benson, S.

Dev. Biol. 148, 261-272, 1991

A/Title: Primary mesenchyme cells of the sea urchin embryo require an autonomously produ

A/Reference number: A43903; PMID:92038439; PMID:1936564

```

A:Accession: A43903
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 7'P, 633-1537, 'G' <WES>
A:Cross-references: GB:S64572; NID:Q238616; PIDN:AA80270.1; PID:Q238617
A:Note: sequence extracted from NCBI backbone (NCBIN:64572, NCBI:P:64573)
R:Venkatesan, M.; De Pablo, F.; Vogeli, G.; Simpson, R.T.
Proc. Natl. Acad. Sci. U.S.A. 83, 3351-3355, 1986
A:Title: Structure and developmentally regulated expression of a Strongylocentrotus purpur
A:Reference number: A23940; MUID:86205894; PMID:3458186
A:Accession: A23940
A:Molecule type: DNA
A:Residues: 742-812 <VEN>
A:Cross-references: EMBL:M13206
C:Superfamily: collagen alpha 1(IV) chain
C:Keywords: coiled coil; extracellular matrix; glycoprotein; trimer; triple helix
F:29-161/Domain: amino-terminal nonhelical, 7S <7SD>
F:162-1523/Region: interrupted helical
F:1524-1752/Domain: carboxyl-terminal nonhelical, NC1 <NC1>
F:1534-1634/Domain: collagen IV carboxyl-terminal repeat <CT1>
F:1644-1748/Domain: collagen IV carboxyl-terminal repeat <CT2>
F:129/Modified site: allysine (lys) #status predicted

Query Match
Best Local Similarity 22.9%; Score 129; DB 2; Length 1752;
Matches 47; Conservative 20; Mismatches 60; Indels 78; Gaps 9;

QY 3 PGGPAAAPORLRLGLLILLILQLPAPSSASEIPKQKQALRRREVVDLYNMCLOGPAGV 62
DB 1453 PGGPPRRDR-----PGPGQ-----PPG-----LTGDKGT 1477
QY 63 PGRDGPAGNGIPGTRGIRGDSFGKGEKCLRESF-----EESMT 103
DB 1478 PGVQGGPGVGVGPEGLKGBQCFKQNGQPGDGGFPGKGEAGLPGSSSGPFTTRHSQT 1537
QY 104 PNYKQC-----SW-----SLANYGIDLGKIAECTFTWRSNSALRYLFGSGLRL 147
DB 1538 TSIPQCPQGTAKMWHYSILFLVQGNRNGHGQDLGKRGSC-----LKRFSMPFLFC-NINN 1592
QY 148 KCRNACCQRRWYFTFGACSGSLPI 172
DB 1593 VCHVASRNDYSYWLSTTE---PMFM 1614

RESULT 3
T20177
Hypothetical protein C53B4.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #ext_change 09-Jul-2004
C:Accession: T20177
R:Berts, M.
Submitted to the EMBL Data Library, December 1995
A:Reference number: Z19233
A:Accession: T20177
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-289 <WIL>
A:Cross-references: UNIPROT:Q18799; EMBL:Z68215; PIDN:CAA92453.1; GSPDB:GN00022; CESP:CS
A:Experimental source: Clone C53B4
A:Genetics:
A:Gene: CESP:C53B4.5
A:Map position: 4

Query Match
Best Local Similarity 9.8%; Score 128; DB 2; Length 289;
Matches 30; Conservative 5; Mismatches 19; Indels 30; Gaps 2;

QY 25 PAPSASSEIPKQKQXQALQREVVVDLYNMCLOGPAGVGRDS-----P 69
DB 203 PDPGSGEGEGRPGQPS-----RGPAGQPKQAGCGPGEKANGEPGP 247
QY 70 GANGIPTGTGIPGRDGFKEKGC 93

```

```

DB 248 GRDGPGRGPRGPRGRHGEKGC 271

RESULT 4
CGHUIS
collagen alpha 1(I) chain precursor - human
N:Alternate names: procollagen alpha 1(I) chain
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1981 #sequence_revision 04-Oct-1996 #ext_change 09-Jul-2004
C:Accession: 160114; S01143; A93335; I55254; A39943; I55237; A35233; S09400; B90567; S11
5269; A29439; I53466; A02852; I37247
R:D'Alessio, M.; Bernard, M.; Pretorius, P.J.; de Wet, W.; Ramirez, F.; Pretorius, P.J.
Gene 67, 105-115, 1988
A:Title: Complete nucleotide sequence of the region encompassing the first twenty-five e
A:Reference number: 160114; MUID:86329734; PMID:284432
A:Accession: 160114
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-369, 'L', 371-589 <DAL>
A:Cross-references: UNIPROT:P02452; UNIPROT:Q14992; UNIPROT:Q16053; UNIPROT:Q13896; UNIP
R:Tromp, G.; Kuivaniemi, H.; Stacey, A.; Shikata, H.; Baldwin, C.T.; Jaenisch, R.; Prock
Biochem. J. 253, 919-922, 1988
A:Title: Structure of a full-length cDNA clone for the prepro-alpha-1(I) chain of human
A:Reference number: S01143; MUID:89025644; PMID:3178743
A:Accession: S01143
A:Molecule type: mRNA
A:Residues: 1-472 <TRO>
A:Cross-references: EMBL:X07884; NID:G30015; PIDN:CAA30731.1; PID:G30016; GB:M36546; NID
A:Note: submitted to the EMBL/GenBank/DBJ databases by Prockop, D.J., 13-JUN-1988
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ding, J.F.; Morabito, M.; Myers, J.; Williams, C.;
Nature 310, 337-340, 1984
A:Title: Human prolalpha1(I) collagen gene structure reveals evolutionary conservation of
A:Reference number: A93335; MUID:84270697; PMID:6462220
A:Accession: A93335
A:Molecule type: DNA
A:Residues: 1-58, 'Q', 60-181 <CHU>
A:Cross-references: EMBL:X00820; NID:G35657; PIDN:CAA25394.1; PID:G35658
R:Rosow, C.M.S.; Vergeer, W.P.; du Plooy, S.J.; Bernard, M.P.; Ramirez, F.; de Wet, W.
J. Biol. Chem. 262, 15151-15157, 1987
A:Title: DNA sequences in the first intron of the human pro-alpha 1(I) collagen gene enh
A:Reference number: I55254; MUID:88033098; PMID:2822714
A:Accession: I55254
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-45 <ROS>
A:Cross-references: GB:J02829; NID:G180387; PIDN:AA51993.1; PID:G180388
R:Bornstein, P.; McKay, J.; Morishima, J.K.; Devarayalu, S.; Gelinas, R.E.
Proc. Natl. Acad. Sci. U.S.A. 84, 8869-8873, 1987
A:Title: Regulatory elements in the first intron contribute to transcriptional control o
A:Reference number: A39943; MUID:88097389; PMID:3480516
A:Accession: A39943
A:Molecule type: DNA
A:Residues: 1-34 <BOR>
A:Cross-references: GB:J03559; NID:G180876; PIDN:AA52052.1; PID:G553238
R:Chu, M.L.; de Wet, W.; Bernard, M.; Ramirez, F.
J. Biol. Chem. 260, 2315-2320, 1985
A:Title: Fine structural analysis of the human pro-alpha 1 (I) collagen gene. Promoter s
A:Reference number: I55237; MUID:85130970; PMID:2857713
A:Accession: I55237
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-34 <CH2>
A:Cross-references: GB:M10627; NID:G180383; PIDN:AA51992.1; PID:G553226
R:Mirtz, W.K.; Keene, D.R.; Hori, H.; Glanville, R.W.; Steinmann, B.; Rao, V.H.; Hollist
J. Biol. Chem. 265, 6312-6317, 1990
A:Title: In vivo and in vitro noncovalent association of excised alpha1(I) amino-termi
A:Reference number: A35233; MUID:90202908; PMID:2318855
A:Accession: A35233
A:Molecule type: protein
A:Residues: 33-52 <WIR>
A:Note: this propeptide fragment remained non-covalently bound to a defective, uncleaved
R:Well, D.; d'Alessio, M.; Ramirez, F.; de Wet, W.G.; Chan, D.; Bateman, J.F.

```

EMBO J. 8, 1705-1710, 1989  
 A>Title: A base substitution in the exon of a collagen gene causes alternative splicing  
 A:Reference number: S09400; MUID:8935664; PMID:2767050  
 A:Accession: S09400  
 A:Molecule type: mRNA  
 A:Residues: 156-183 <ME>  
 R:Click, E.M.; Bornstein, P.  
 Biochemistry 9, 4699-4706, 1970  
 A>Title: Isolation and characterization of the cyanogen bromide peptides from the alpha1  
 A:Reference number: A90567; MUID:71038625; PMID:5529814  
 A:Accession: CNBR0-1, CNBR2, CNBR4, CNBR5  
 A:Molecule type: protein  
 A:Residues: 162-198, 'Z', 200-201, 'Z', 203-206, 'Z', 208-209, 'Z', 211-228, 'B', 230, 'BB', 233, 'Z'  
 A:Experimental source: skin  
 A>Note: evidence for 170-alanine  
 R:Beyle, B.; Nohrbom, H.; Diebold, J.; Lehmann, H.; Bodo, M.; Deutzmann, R.; Mueller, F.  
 Eur. J. Biochem. 192, 153-159, 1990  
 A>Title: A critical crosslink region in human bone-derived collagen type I. Specific cle  
 A:Reference number: S11372; MUID:90382436; PMID:2169412  
 A:Accession: S11372  
 A:Molecule type: protein  
 A:Residues: 175-187, 274-287, 'P', 289 <BAE>  
 A>Note: sequence of collagen alpha 1(I) isolated from bone after pepsin digestion  
 R:Peak, S.B.; Scholz, P.M.; Amenta, P.S.; Constantinou, C.D.; Levi-Mizli, S.A.; Gonzalez  
 J. Biol. Chem. 266, 21827-21837, 1991  
 A>Title: The substitution of arginine for glycine 85 of the alpha 1(I) procollagen chain  
 A:Reference number: I55342; MUID:92042092; PMID:1718984  
 A:Accession: I55342  
 A:Molecule type: mRNA  
 A:Status: translated from GB/EMBL/DBJ  
 A:Residues: 258-268, 1347-1357 <DE>  
 A:Cross-references: GB:S6795; NID:G239007; PIDN:AA02350.1; PID:G239008  
 A>Note: sequences from the 5' and 3' ends only are shown; mutant sequence 263-Arg report  
 R:Morgan, P.H.; Jacobs, H.G.; Segrest, J.P.; Cunningham, L.W.  
 J. Biol. Chem. 245, 5042-5048, 1970  
 A>Title: Comparative study of glycoproteins derived from selected vertebrate collagens.  
 A:Reference number: A92069; MUID:11001508; PMID:4319110  
 A:Accession: A92069  
 A:Molecule type: protein  
 A:Residues: 263-268 <MOR>  
 A:Experimental source: skin  
 A>Note: attachment of 2-O-alpha-D-glucosyl-O-beta-D-galactose to 5-hydroxylysine  
 R:Labhard, M.E.; Hollister, D.W.  
 Matrix 10, 124-130, 1990  
 A>Title: Segmental amplification of the entire helical and telopeptide regions of the CD  
 A:Reference number: S15989; MUID:90326017; PMID:2374517  
 A:Accession: S15989  
 A:Molecule type: mRNA  
 A:Residues: 281-302, 402-420, 823-843, 925-944, 1026-1045, 1143-1162 <LAB>  
 R:Wirtz, M.K.; Rao, V.H.; Glanville, R.W.; Labhard, M.E.; Pretorius, P.J.; de Vries, W.N.  
 Connect. Tissue Res. 29, 1-11, 1993  
 A>Title: A cysteine for glycine substitution at position 175 in an alpha 1 (I) chain of  
 A:Reference number: I52905; MUID:93339042; PMID:8339541  
 A:Accession: I52905  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 342-352, 'C', 354-359 <MI2>  
 A:Cross-references: GB:S64717; NID:G408195; PIDN:AA02767.1; PID:G408196  
 A>Note: mutant sequence from patient with osteogenesis imperfecta  
 R:Bernard, M.P.; Chu, M.L.; Wiers, J.C.; Ramirez, F.; Eikenberry, E.F.; Prockop, D.J.  
 Biochemistry 22, 5213-5223, 1983  
 A>Title: Nucleotide sequences of complementary deoxyribonucleic acids for the proalpha1  
 A:Reference number: A90476; MUID:84080385; PMID:6689127  
 A:Accession: A90476  
 A:Molecule type: mRNA  
 A:Residues: 425-1128, 'X', 1252-1328, 'S', 1330-1390, 'X', 1397-1464 <BR>  
 A:Cross-references: GB:X01228; NID:G180391; PIDN:AA051995.1; PID:G180392  
 A>Note: sequence partially completed for missing nucleotides by A29439  
 R:Chu, M.L.; Garayulu, V.; Williams, C.J.; Ramirez, F.  
 J. Biol. Chem. 260, 691-694, 1985  
 A>Title: Multixon deletion in an osteogenesis imperfecta variant with increased type I

A:Reference number: A22161; MUID:85104934; PMID:2981843  
 A:Accession: A22161  
 A:Molecule type: DNA  
 A:Residues: 472-594, 'R', 596-607 <CH3>  
 A:Cross-references: GB:X03178; GB:X03179; NID:G179612; NID:G179613; PIDN:AA051847.1; PI  
 A>Note: the authors translated the codon CGT for residue 595 as Pro  
 R:Walls, G.A.; Starman, B.J.; Zimm, A.B.; Byers, P.H.  
 Am. J. Hum. Genet. 46, 1034-1040, 1990  
 A>Title: Variable expression of osteogenesis imperfecta in a nuclear family is explained  
 A:Reference number: A35336; MUID:90252792; PMID:2339700  
 A:Accession: A35336  
 A:Molecule type: mRNA  
 A:Residues: 710-720, 'E', 722-737, 'E', 739-745 <MAL>  
 A>Note: the authors translated the codons CAG for 721 and CGT for 738 as Glu  
 R:Forlino, A.; Zolezzi, F.; Valli, M.; Pignatelli, P.F.; Cetta, G.; Brunelli, P.C.; Motte  
 Hum. Mol. Genet. 3, 2201-2206, 1994  
 A>Title: Severe (type III) osteogenesis imperfecta due to glycine substitutions in the  
 A:Reference number: I54365; MUID:95187161; PMID:7881420  
 A:Accession: I54365  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 746-766, 'S', 768-781 <FOR>  
 A:Cross-references: GB:I47667; NID:G1009093; PIDN:AA059576.1; PID:G1009094  
 R:Chesler, S.D.; Wallis, G.A.; Byers, P.H.  
 J. Biol. Chem. 268, 18218-18225, 1993  
 A>Title: Mutations in the carboxyl-terminal propeptide of the pro alpha 1(I) chain of ty  
 A:Reference number: A47426; MUID:93352646; PMID:8349697  
 A:Accession: A47426  
 A:Molecule type: mRNA  
 A:Residues: 1179-1276, 'H', 1278-1336, 1339-1387, 'R', 1389-1464 <CH>  
 A:Cross-references: GB:S64596; NID:G407589; PIDN:AA057856.1; PID:G407590  
 A>Note: sequence extracted from NCBI backbone (NCBI:136444, NCBI:136445)  
 A:Accession: B47426  
 A:Molecule type: mRNA  
 A:Residues: 1179-1464 <CH4>  
 A:Experimental source: normal dermal fibroblast culture  
 A:Accession: C47426  
 A:Molecule type: mRNA  
 A:Residues: 1179-1276, 'H', 1278-1464 <CH5>  
 A:Experimental source: fetal cell 86-237  
 A:Accession: D47426  
 A:Molecule type: mRNA  
 A:Residues: 1179-1336, 1339-1464 <CH6>  
 A:Experimental source: fetal cell 86-146  
 A:Accession: E47426  
 A:Molecule type: mRNA  
 A:Residues: 1179-1387, 'R', 1389-1464 <CH7>  
 A:Experimental source: fetal cell 88-251  
 R:Chou, D.H.; Apone, S.; Byre, D.R.; Starman, B.J.; Andraessen, P.; Charbonneau, H.; Nic  
 J. Biol. Chem. 263, 14605-14607, 1988  
 A>Title: Substitution of Cysteine for Glycine within the Carboxyl-terminal Telopeptide c  
 A:Reference number: I55269; MUID:89008319; PMID:3170557  
 A:Accession: I55269  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1187-1194, 'C', 1196-1220 <COH>  
 A:Cross-references: GB:M23213; NID:G340842; PIDN:AA059363.1; PID:G499622  
 A>Note: mutant sequence from a patient with mild osteogenesis imperfecta  
 R:Maehle, J.K.; Raasing, M.; Vitra, A.; Vuorio, E.  
 Nucleic Acids Res. 16, 349, 1988  
 A>Title: Human pro-alpha-1(I) collagen: cDNA sequence for the C-propeptide domain.

Query Match 9.8%; Score 127.5; DB 1; Length 1464;  
 Best Local Similarity 37.4%; Pred. No. 0.0045;  
 Matches 34; Conservative 2; Mismatches 22; Indels 33; Gaps 4;  
 3 PGGPAAAS--PQRLGILLILLILQLPAPSSASRTPKCKAKAQRGRVVDVLYMGKLCGPA 60  
 DB 357 PGGPAAAS--PQRLGILLILLILQLPAPSSASRTPKCKAKAQRGRVVDVLYMGKLCGPA 385  
 61 GVGGRGSPGANGIPOTPIGPDGFKGEKG 91



A:Residues: 1427-1444 <BPR>  
A:Note: sequence extracted from NCBI backbone (NCBIP:133363): sequence incorrectly ident  
R:Morrison, K.E.; Matiyama, M.; Yang-Feng, T.L.; Redders, S.T.  
Am. J. Hum. Genet. 49, 545-554, 1991  
A:Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of  
A:Reference number: A39786; MUID:91353570; PMID:1882840  
A:Accession: A39786  
A:Molecule type: mRNA  
A:Residues: 1453-1593, 'A', 1595-1670 <NOR>  
A:Cross-references: GB:555790; NID:g224418; PIDN:AA819637.1; PID:g224419  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope  
C:Genetics:  
A:Gene: COL4A3  
A:Cross-references: GDB:128351; OMIM:120070  
A:Map position: 2q36-2q37  
A:Intons: 1385/1, 1418/1, 1488/1, 1547/2, 1585/3, 1643/2 #status incomplete  
A:Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3  
among times: amino-terminal domains (with disulfide and desmosine cross-links), dimeric  
er associations in the interrupted helical domain (with disulfide and desmosine cross-  
C:Function:  
A:Description: minor structural component of extracellular basement membrane in kidney  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracel  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAT>  
F:29-42/Domain: amino-terminal nonhelical, NH1 <NHI>  
F:43-1438/Region: interrupted helical  
F:791-793/Region: cell attachment (R-G-D) motif  
F:966-998/Region: cell attachment (R-G-D) motif  
F:1154-1156/Region: cell attachment (R-G-D) motif  
F:1306-1308/Region: cell attachment (R-G-D) motif  
F:1345-1347/Region: cell attachment (R-G-D) motif  
F:1432-1434/Region: cell attachment (R-G-D) motif  
F:1439-1670/Domain: carboxyl-terminal nonhelical, NC1 <NC1>  
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:31,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status predi  
F:253/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted  
F:1505-1511,1616-1622/Disulfide bonds: #status predicted  
F:1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted  
Query Match 9.6%; Score 125; DB 1; Length 1670;  
Best Local Similarity 36.2%; Pred. No. 0.0086;  
Matches 34; Conservative 6; Mismatches 36; Indels 18; Gaps 3;  
QY 5 GPAASPRQLRGLLLLLQLPAPSSASEIRKQKQALRQREVVDLYNGMCLGGPAGVPG 64  
DB 467 GPPGEPR---GLLCTOCPTVTPGPGPLGLGLHGVKQIPR-----GGAAGLKG 511  
QY 65 RDGSPGANGIPGTPTGIPGRD---GFKGEKECELR 95  
DB 512 SPGSPGNTGIPGFPFGAGAGDGLKGEKEETIQ 545  
RESULT 9  
CGHUTL  
collagen alpha 1(III) chain precursor - human  
N:Alternate names: procollagen alpha 1(III) chain  
C:Species: Homo sapiens (man)  
C:Date: 24-Apr-1984 #sequence revision 01-Sep-1995 #text change 09-Jul-2004  
C:Accession: S05272; S04642; PE0011; S01726; S04887; A90399; A94562; I51868; S59511; A90  
R:Prockop, D.J.  
submitted to the EMBL Data Library, February 1989  
A:Reference number: S05272  
A:Accession: S05272  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1240, 'V', 1242-1466 <PRC>  
A:Cross-references: UNIPROT:P02461; EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g30058

R:Ala-Kokko, L.; Kontunauri, S.; Baldwin, C.T.; Kuivaniemi, H.; Prockop, D.J.  
Biochem. J. 260, 509-516, 1989  
A:Title: Structure of cDNA clones coding for the entire prepro-alpha1(III) chain of hum  
ences.  
A:Reference number: S04642; MUID:89350838; PMID:2764886  
A:Accession: S04642  
A:Molecule type: mRNA  
A:Residues: 1-1196 <ALA>  
A:Cross-references: EMBL:X14420; NID:g30057; PIDN:CAA32583.1; PID:g30058  
A:Note: the complete sequence is not shown  
R:Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.  
Gene 78, 255-265, 1989  
A:Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene  
A:Reference number: PE0011; MUID:89378752; PMID:2777083  
A:Accession: PE0011  
A:Molecule type: DNA  
A:Residues: 1-176 <BEN>  
A:Cross-references: GB:M26939; NID:g180813; PIDN:AA52040.1; PID:g180814  
R:Roman, P.D.; Ricci, G.A.; de Crombrughe, B.  
Nucleic Acids Res. 16, 7201, 1988  
A:Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pr  
A:Reference number: S01726; MUID:88303360; PMID:3405773  
A:Accession: S01726  
A:Molecule type: mRNA  
A:Residues: 1-170 <TOM>  
A:Cross-references: EMBL:X07240; NID:g30060; PIDN:CAA30229.1; PID:g30061  
A:Note: the authors translated the codon CAG for residue 154 as His  
R:Janeczko, R.A.; Ramirez, F.  
Nucleic Acids Res. 17, 6742, 1989  
A:Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen  
A:Reference number: S04887; MUID:89386015; PMID:2780304  
A:Accession: S04887  
A:Molecule type: mRNA  
A:Residues: 149-163, 'G', 164-240, 'D', 242-471, 'D', 473-487, 'L', 489, 'S', 491-613, 'Y', 615-634  
A:Cross-references: EMBL:X15332; NID:g29545; PIDN:CAA3387.1; PID:g930045  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 16, 1158-1164, 1977  
A:Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptid  
A:Reference number: A90399; MUID:77134724; PMID:557355  
A:Accession: A90399  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-232, 'P', 234-292, 'D', 294-398 <SEY1>  
A:Experimental source: liver  
A:Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galact  
R:Sever, J.M.  
submitted to the Atlas, December 1977  
A:Reference number: A94562  
A:Accession: A94562  
A:Molecule type: protein  
A:Residues: 'V', 169-225, 229-277, 'A', 279-292, 'D', 294, 'S', 296-398 <SEY2>  
A:Experimental source: liver  
A:Note: author submitted corrections to A90399  
R:Milwicz, D.M.; Wlitz, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.  
Am. J. Hum. Genet. 53, 62-70, 1993  
A:Title: Parental somatic and germ-line mosaicism for a multiexon deletion with unusual  
leping.  
A:Reference number: I51868; MUID:93304430; PMID:8317500  
A:Accession: I51868  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 186-194 <MTL>  
A:Cross-references: GB:S62925; NID:g386425; PIDN:AAD13937.1; PID:g4261637  
R:Chiodo, A.A.; Silence, D.O.; Cole, W.G.; Bateman, J.F.  
Biochem. J. 311, 939-943, 1995  
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3  
A:Reference number: S59511; MUID:96067614; PMID:7487954  
A:Accession: S59511  
A:Molecule type: mRNA  
A:Residues: 302-423 <CHI>  
A:Cross-references: GB:S79877; NID:g1195576; PIDN:AA835615.1; PID:g1195577  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 17, 3404-3411, 1978



A>Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBr p  
A:Reference number: A90414; PMID:79000343; PMID:667591  
A:Accession: A90414  
A:Molecule type: protein  
A:Residues: 399-675, 'N', 677-727 <SEY3>  
A:Experimental source: liver  
R:Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
J. Biol. Chem. 266, 5256-5259, 1991  
A>Title: G to T transversion at position +5 of a splice donor site causes skipping of th  
A:Reference number: I55349; PMID:91161621; PMID:1672129  
A:Accession: I55349  
A:Molecule type: DNA  
A:Status: translated from GB/EMBL/DBJ  
A:Residues: 537-605 <LBE>  
A:Cross-references: GB:M59312; NID:g180815; PIDN:AAA52041.1; PID:g180816  
R:Sever, J.M.; Mainardi, C.; Kang, A.H.  
Biochemistry 19, 1583-1589, 1980  
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB5 from ty  
A:Reference number: A90438; PMID:80198282; PMID:6246925  
A:Accession: A90438  
A:Molecule type: protein  
A:Residues: 728-895, 'A', 897-964 <SEY4>  
A:Experimental source: liver  
R:Cole, W.G.; Chiodo, A.A.; Lemande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Char  
J. Biol. Chem. 265, 17070-17077, 1990  
A>Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping an  
A:Reference number: A38303; PMID:91009133; PMID:2145268  
A:Accession: A38303  
A:Molecule type: mRNA  
A:Residues: 861-1015 <COL>  
A:Cross-references: GB:j05617; GB:M55603; GB:M59227; NID:g180878; PIDN:AA59383.1; PID:g  
A>Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syn  
Nucleic Acids Res. 16, 2337, 1988  
A>Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.  
A:Reference number: S02119; PMID:88189827; PMID:3357782  
A:Accession: S02119  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 950-1018, 'V', 1020-1183, 'S', 1185-1466 <MAN>  
A:Cross-references: EMBL:X06700; NID:g30053; PIDN:CAA29886.1; PID:g30054  
R:Sever, J.M.; Kang, A.H.  
Biochemistry 20, 2621-2627, 1981  
A>Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB9 from ty  
A:Reference number: A90446; PMID:81208139; PMID:7016180  
A:Accession: A90446  
A:Molecule type: protein  
A:Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-  
A:Experimental source: liver  
R:Loide, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Mye  
Nucleic Acids Res. 12, 9383-9394, 1984  
A>Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollagen  
A:Reference number: A93551; PMID:85087944; PMID:6096827  
A:Accession: A93551  
A:Molecule type: mRNA  
A:Residues: 1065-1155, 'P', 1157-1466 <LOI>  
A:Cross-references: EMBL:X01655; EMBL:X01742; NID:g29584; PIDN:CAA25821.1  
R:Wiskuln, M.; Dalgleish, R.; Kluge-Beckerman, B.; Renard, S.I.; Tolstoshev, P.; Brant  
Biochemistry 23, 1408-1413, 1986  
A>Title: Human type III collagen gene expression is coordinately modulated with the type  
A:Reference number: I52393; PMID:86187804; PMID:3754462  
A:Accession: I52393  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1161-1200 <MTS>  
A:Cross-references: GB:M13146; NID:g180415; PIDN:AAA52003.1; PID:g180416  
R:Emanuel, B.S.; Camizazaro, L.A.; Sever, J.M.; Myers, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985  
A>Title: Human alpha 1(III) and alpha 2(IV) procollagen genes are located on the long arm  
A:Reference number: I59025; PMID:85216505; PMID:3658826  
A:Accession: I79359  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Residues: 1165-1196 <EMA>  
A:Cross-references: GB:M11134; NID:g180417; PIDN:AAA52004.1; PID:g180418  
R:Chu, M.L.; Weil, D.; de Wet, W.; Bernard, M.; Sipola, M.; Ramirez, F.  
J. Biol. Chem. 260, 4357-4363, 1985  
A>Title: Isolation of cDNA and genomic clones encoding human pro-alpha1(III) collagen. P  
A:Reference number: A92516; PMID:85157600; PMID:2579949  
A:Accession: A92516  
A:Molecule type: DNA  
A:Residues: 1176-1240, 'V', 1242-1356, 'P', 1358-1466 <CHU>  
A:Cross-references: GB:M10613; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB  
A:Experimental source: liver  
A>Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given f  
action  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit ('  
3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently O  
C:Genetics:  
A:Gene: GDB:COL3A1  
A:Cross-references: GDB:118729; OMIM:120180  
A:Map position: 2q31-2q31  
A:Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1275/1; 1337/3; 1418/3  
A>Note: the list of introns is incomplete; defects in this gene can result in Ehlers-Dan  
C:Complex: type III collagen is a homotrimer of monomers initially linked by disulfide b  
er of their length, is formed with desmosine cross-links made from lysine and allysine r  
C:Function:  
A:Description: structural component of extracellular fibrous polymer that maintains inte  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
C:Keywords: coiled coil, Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hyd  
F:1-23/Domain: signal sequence #status predicted <SIG>  
F:24-153/Domain: amino-terminal propeptide #status predicted <PRO>  
F:31-91/Domain: von Willebrand factor type C repeat homology <WVC>  
F:154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>  
F:154-167/Region: amino-terminal nonhelical telopeptide  
F:168-1196/Region: helical  
F:1091-1093/Region: cell attachment (R-G-D) motif  
F:1197-1221/Region: carboxyl-terminal nonhelical telopeptide  
F:1222-1466/Domain: carboxyl-terminal propeptide #status predicted <CPR>  
F:1238-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
F:24/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted  
F:153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted  
F:154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicte  
F:154/Modified site: allysine (Lys) #status predicted  
F:653,284,860,977,1106/Modified site: 5-hydroxylysine (Lys) #status experimental  
F:263/Binding site: carboxylate (Lys) (covalent) #status experimental  
F:584,1094/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental  
F:948-949/Cleavage site: Gly-Ile (collagenase) #status experimental  
F:1106/Binding site: carboxylate (Lys) (covalent) #status predicted  
  
Query Match 9.5%; Score 124; DB 1; Length 1466;  
Best Local Similarity 37.0%; Pred. No. 0.0092;  
Matches 40; Conservative 8; Mismatches 46; Indels 14; Gaps 5;  
  
QY 22 LQLPAPSSASRPPKQKQAKQRLQREVVLY--NGWCLGPGVPGDPSPGANGIPGPG 79  
DB 72 LDCENP---LTPFGCCAVCPDPPTAPTRPPNGGPGKDPDPGIPGNGDPIRG 127  
  
QY 80 IPRGPGFKGKGECGRSEF---EESGTPNYKQCSSTNATYGLDKRIA 124  
DB 128 QPGSGSGSPGPGIC--ESCPGPGQNYSPQYSDYDKS---GVAVGCLA 170  
  
RESULT 10  
S16366  
collagen alpha 2(IV) chain precursor - pig roundworm  
C:Species: Acanthis suum (pig roundworm)  
C:Date: 04-Dec-1992 #sequence\_revision 04-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: S16366  
R:Pettilt, J.; Kingston, I.B.  
J. Biol. Chem. 266, 16149-16156, 1991  
A>Title: The complete primary structure of a nematode alpha-2(IV) collagen and the parti  
A:Reference number: S16366; PMID:91340768; PMID:1771907  
A:Accession: S16366  
A:Molecule type: mRNA  
A:Residues: 1-1763 <JBI>



A:Cross-references: UNIPROT:P27393; GB:M67507; NID:g159648; PIDN:AAA18014.1; PID:g159648  
 C:Genetics:  
 A:Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;  
 C:Superfamily: collagen alpha 1(IV) chain  
 C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfid  
 F:1-26/Domain: signal sequence #status predicted <SIG>  
 F:27-1763/Domain: collagen alpha 2(IV) chain #status predicted <MNT>  
 F:27-42/Domain: non-collagenous N1 #status predicted <NHL>  
 F:43-1529/Domain: collagenous #status predicted <COL>  
 F:197-199/Region: cell attachment (R-G-D) motif  
 F:1530-1763/Domain: carboxyl-terminal nonhelical, NCI #status predicted <NC1>  
 F:1530-1658/Domain: repeat NCI #status predicted <NC11>  
 F:1539-1763/Domain: repeat NCI #status predicted <NC12>  
 F:31,34,39,41,536,539/disulfide bonds: interchain #status predicted  
 F:126/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:1593-1599;1702-1709/disulfide bonds: #status predicted

Query Match 9.5%; Score 124; DB 2; Length 1763;

Best Local Similarity 25.8%; Pred. No. 0.011;

Matches 67; Conservative 24; Mismatches 93; Indels 76; Gaps 16;

QY 5 GPAPSPQRLRLGLLILLLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQGPAGV 62

DB 1449 GPWGAPEGRGEGKGLPGLDGLPGPSG-----PGFPAKGR-----DGFPGQPMPSKGA 1498

QY 63 PGRDSPGANGIGTGTGIRGDRGPKGKGLRESF---ESWTPTNKQCS-----W-- 111

DB 1499 PGLPFPFGLRGIRGDRGPKGKGLRESF---ESWTPTNKQCS-----W-- 111

QY 112 -----SLANGYDIDGKIAECTFTMRGNSALRVLSGLRLKGNAC-----CQR 156

DB 1559 YSLIYTBGKNSHNOULGHAGSC-----LSRFTMPPLF-----CDVNVNNTVASRNDKSY 1609

QY 157 WYTFNGACSGPPLPIEALIIYLDQSGPEMNSTINIRHTSSVEGLGIGALVDV---AI 213

DB 1610 WLST-----TADIPMPV-----SEGIEPIYSIRCAVCEA-PANVIIVHSQRT 1651

QY 214 WVGTCSDYPRKGDASTGMSV 233

DB 1652 QIPNCPN-----GWNSTL 1663

## RESULT 11

T15268

hypothetical protein F59E12.12 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C/Accession: T15268

R:Johnson, D.

submitted to the EMBL Data Library, May 1997

A&gt;Description: The sequence of C. elegans cosmid F59E12.

A/Reference number: Z18318

A/Accession: T15268

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-311 &lt;JON&gt;

A:Cross-references: UNIPROT:O01904; EMBL:AF003386; NID:g2088833; PID:g2088834; PIDN:AA5

A/Experimental source: strain Bristol N2; clone F59E12

C:Genetics:

A:Gene: CESP:F59E12.12

A:Map position: 2

A:Introns: 24/2

Query Match 9.5%; Score 123.5; DB 2; Length 311;

Best Local Similarity 53.3%; Pred. No. 0.0018;

Matches 24; Conservative 4; Mismatches 12; Indels 5; Gaps 1;

QY 53 GWC-----LOGPAGVGRDSSPGANGIPGTGIRGDRGPKGKGE 92

DB 227 GKCDENVNVAGQPPGSPGPPGLPDPDGLPGPAGNPGDGGGSPAGE 271

## RESULT 12

## T26185

hypothetical protein W05B2.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T26185

R:Gardner, A.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z20166

A/Accession: T26185

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-304 &lt;WIL&gt;

A:Cross-references: UNIPROT:Q9Y349; EMBL:Z81138; PIDN:CAH03475.1; GSPDB:GN00021; CESP:W

A/Experimental source: clone W05B2

C:Genetics:

A:Gene: CESP:W05B2.1

A:Map position: 3

A:Introns: 27/3

Query Match 9.4%; Score 123; DB 2; Length 304;

Best Local Similarity 33.0%; Pred. No. 0.0019;

Matches 30; Conservative 5; Mismatches 24; Indels 32; Gaps 2;

QY 3 PGAPASPQRLRLGLLILLLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQGPAGV 62

DB 227 PQGP-----PGAPSPGARGGPGQAG-----APGPKGP 254

QY 63 PGRDSPGANGIPGTGIRGDRGPKGKGE 93

DB 255 SGAPGQPGADGNPAGPQGGGAGGEGKIC 285

## RESULT 13

T26184

hypothetical protein W05B2.6 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C&gt;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C/Accession: T26184

R:Gardner, A.

submitted to the EMBL Data Library, October 1996

A/Reference number: Z20166

A/Accession: T26184

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-304 &lt;WIL&gt;

A:Cross-references: UNIPROT:Q9YV3; EMBL:Z81138; PIDN:CAH03474.1; GSPDB:GN00021; CESP:W

A/Experimental source: clone W05B2

C:Genetics:

A:Gene: CESP:W05B2.6

A:Map position: 3

A:Introns: 27/3

Query Match 9.4%; Score 123; DB 2; Length 304;

Best Local Similarity 33.0%; Pred. No. 0.0019;

Matches 30; Conservative 5; Mismatches 24; Indels 32; Gaps 2;

QY 3 PGAPASPQRLRLGLLILLLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQGPAGV 62

DB 227 PQGP-----PGAPSPGARGGPGQAG-----APGPKGP 254

QY 63 PGRDSPGANGIPGTGIRGDRGPKGKGE 93

DB 255 SGAPGQPGADGNPAGPQGGGAGGEGKIC 285

## RESULT 14

S22917

collagen alpha 5(IV) chain precursor, renal splice form - human

N/Alternate names: procollagen alpha 5(IV) chain

N/Contains: collagen alpha 5(IV) chain precursor, leukocyte splice form

C:Species: Homo sapiens (man)

C&gt;Date: 30-Sep-1993 #sequence\_revision 27-Feb-1997 #text\_change 09-Jul-2004

C/Accession: S22917; A54365; A57079; A37122; I54317; A34850; S18850; I56971; I76598; A35

R.Zhou, J.; Hertz, J.M.; Leinonen, A.; Trygsvaen, K.  
J. Biol. Chem. 267, 12475-12481, 1992  
A:Title: Complete amino acid sequence of the human alpha-5(IV) collagen chain and identification of the Alport syndrome patient.  
A:Reference number: S22917; PMID:92316923; PMID:1352287  
A:Accession: S22917  
A:Molecule type: mRNA  
A:Residues: 1-967 <ZHO>  
A:Cross-references: UNIPROT:P29400; GB:M90464; NID:g180826; PIND:AAA52046.1; PID:g553234  
R.Zhou, J.; Leinonen, A.; Trygsvaen, K.  
J. Biol. Chem. 269, 6608-6614, 1994  
A:Title: Structure of the human type IV collagen COL4A5 gene.  
A:Reference number: A54365; PMID:94165049; PMID:8120014  
A:Accession: A54365  
A:Molecule type: DNA  
A:Residues: 1-922 <ZHZ>  
A:Cross-references: GB:U04470; NID:g463378; GB:U04520; NID:g463428; PIND:AA027816.1; PID:R.Zhou, J.; Mochizuki, T.; Smeets, H.; Antignac, C.; Laurila, P.; de Paep, A.; Trygsvaen Science 261, 1167-1169, 1993  
A:Title: Deletion of the paired alpha5(IV) and alpha6(IV) collagen genes in inherited sm  
A:Reference number: A57079; PMID:93361972; PMID:8356449  
A:Accession: A57079  
A:Molecule type: DNA  
A:Residues: 1-27 <ZHD>  
A:Cross-references: GB:Z37153; NID:g587203; PIND:CAA6512.1; PID:g587204  
R.Pihlajaniemi, T.; Pohlajainen, E.R.; Myer, J.C.  
J. Biol. Chem. 265, 13758-13766, 1990  
A:Title: Complete primary structure of the triple-helical region and the carboxyl-terminal  
A:Reference number: A37122; PMID:90337990; PMID:2380186  
A:Accession: A37122  
A:Molecule type: mRNA  
A:Residues: 84-439, 'GS', 442-624, 'LALG', 629-666, 'FR', 669-887, 'R', 889-1264, 1271-1691 <PIH>  
A:Cross-references: GB:U05558; EMBL:M58526; NID:g3134209  
A:Note: Submitted to the EMBL Data Library, February 1991  
A:Note: the authors translated the codon GCC for residue 115 as Val  
R.Rendert, A.; Serl, M.; Myer, J.C.; Pihlajaniemi, T.; Massella, L.; Rizzoni, G.; De Ma  
Hum. Mol. Genet. 1, 127-129, 1992  
A:Title: De novo mutation in the COL4A5 gene converting glycine 325 to glutamic acid in  
A:Reference number: I54317; PMID:93244772; PMID:1363780  
A:Accession: I54317  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 313-324, 'E', 326-330 <REN>  
A:Cross-references: GB:G59934; NID:g299946; PIND:AA013909.1; PID:g4261609  
R.Hoslika, S.L.; Bady, R.L.; Byers, M.G.; Hoehly, M.; Shows, T.B.; Trygsvaen, K.  
Proc. Natl. Acad. Sci. U.S.A. 87, 1606-1610, 1990  
A:Title: Identification of a distinct type IV collagen alpha chain with restricted kidney  
A:Reference number: A34850; PMID:90160375; PMID:1689491  
A:Accession: A34850  
A:Molecule type: mRNA  
A:Residues: 914-1264, 1271-1691 <HOS>  
A:Cross-references: EMBL:M1115; NID:g180824; PIND:AAA52045.1; PID:g180825  
R.Zhou, J.; Hoslika, S.L.; Chow, L.T.; Trygsvaen, K.  
Genomics 9, 1-9, 1991  
A:Title: Characterization of the 3' half of the human type IV collagen alpha-5 gene that  
A:Reference number: A37969; PMID:91169491; PMID:2004755  
A:Accession: A37969  
A:Molecule type: DNA  
A:Residues: 924-1264, 1271-1691 <ZHZ>  
A:Cross-references: EMBL:M63455; EMBL:M63456; EMBL:M63457; EMBL:M63458; EMBL:M63459; EMBL:M63460; EMBL:M63471; EMBL:M63472; EMBL:M63473; NID:g177922; PIND:AAA51558.1; PID:R.Zhou, J.; Van Damme, B.; Van Damme-Lombaerts, R.; Van den Berghe, H.; Cassiman, J.J.; M  
Kidney Int. 44, 1316-1321, 1993  
A:Title: Differential splicing of COL4A5 mRNA in kidney and white blood cells: a complex  
A:Reference number: I56971; PMID:9413540; PMID:8301933  
A:Accession: I56971  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1258-1276 <GUO>  
A:Cross-references: GB:869168; NID:g545095; PIND:AA06012.1; PID:g545096  
A:Note: Kidney splice form  
A:Accession: I76598  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1284-1291, 'TRIGLYACTV' <GUO2>  
A:Cross-references: GB:S69169; NID:g545097; PIND:AA06013.1; PID:g545098  
A:Note: frameshift mutation in patient with Alport syndrome  
R.Myers, J.C.; Jones, T.A.; Pohlajainen, E.R.; Kadi, A.S.; Goddard, A.D.; Sheer, D.; So  
Am. J. Hum. Genet. 46, 1024-1033, 1990  
A:Title: Molecular cloning of alpha5(IV) collagen and assignment of the gene to the regi  
A:Reference number: A35335; PMID:90252791; PMID:2339699  
A:Accession: A35335  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1448-1477 <MYE>  
R.Nakazato, H.; Hattori, S.; Ushijima, T.; Matsura, T.; Koitabashi, Y.; Takada, T.; Yos  
Kidney Int. 46, 1307-1314, 1994  
A:Title: Mutations in the COL4A5 gene in Alport syndrome: a possible mutation in primord  
A:Reference number: I56975; PMID:95156993; PMID:7853788  
A:Accession: I56975  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1595-1602 <NMK>  
A:Cross-references: GB:S75903; NID:g913882; PIND:AA033374.1; PID:g913883  
A:Note: premature termination mutation from a patient with Alport syndrome; one other mu  
R.Lemink, H.H.; Schroeder, C.H.; Brunner, H.G.; Nelen, M.R.; Zhou, J.; Trygsvaen, K.;  
Genomics 17, 485-489, 1993  
A:Title: Identification of four novel mutations in the COL4A5 gene of patients with Alpo  
A:Reference number: I54188; PMID:94010948; PMID:8406498  
A:Accession: I54188  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1604-1607, 'VHDVAVK' <LEM>  
A:Cross-references: GB:S65767; NID:g425563; PIND:AA013967.1; PID:g4261667  
A:Note: frameshift mutation from a patient with Alport syndrome; five other mutations ar  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit ('  
ed and subsequently O-glycosylated).  
C:Genetics:  
A:Gene: GDB:COL4A5; AFS  
A:Cross-references: GDB:120596; OMIM:303630  
A:Map position: Xq22-Xq22  
A:Intons: 27/3; 47/3; 77/3; 92/3; 107/3; 128/3; 146/3; 155/3; 182/3; 203/3; 215/3; 229/  
3; 799/1; 837/1; 893/1; 923/1; 973/1; 1006/1; 1036/1; 1082/3; 1125/1; 1152/1; 1155/1; 1  
A:Note: the alpha 5(IV) and alpha 6(IV) chain genes are encoded on opposite strands with  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 5(  
mory trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-l  
C:Function:  
A:Description: minor structural component of extracellular basement membrane  
A:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: Alport syndrome; basement membrane; coiled coil; extracellular matrix; glyco  
F:1-26/Domain: signal sequence #status predicted <SIG>  
F:27-1691/Product: collagen alpha 5(IV) chain, renal splice form #status predicted <MAT1  
F:127-1264, 1271-1691/Product: collagen alpha 5(IV) chain, leukocyte splice form #status p  
F:77-41/Domain: amino-terminal nonhelical, NC2 #status predicted <NC2>  
F:12-1462/Region: interrupted helical  
F:1463-1691/Domain: interrupted helical nonhelical, NC1 #status predicted <NC1>  
F:1473-1573/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1583-1687/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:29.32.38.40.124.451.481.484/Disulfide bonds: interchain #status predicted  
F:125/Binding site: carbohydrate (Aan) (covalent) #status predicted  
F:1482-1570, 1515-1573/Disulfide bonds: (or 1482-1573, 1515-1570) #status predicted  
F:1529-1533, 1638-1644/Disulfide bonds: #status predicted  
F:1592-1684, 1626-1687/Disulfide bonds: (or 1592-1687, 1626-1684) #status predicted

Query Match 94% Score 123; DB 1; Length 1691;  
Best local Similarity 36.5%; Pred. No. 0.013; Indels 16; Gaps 4;  
Matches 35; Conservative 7; Mismatches 38;  
QY 3 PGAPAA--SPQRLRGILLILLILQLPAPSSASERP-----KGKQKALQRREVVLDLYNQC 55  
DB 1256 PGCGPRGPRPTGFG-----LPGCGPRGRLPENGKIKG-EKKNPPQGLPGLGLK 1306  
QY 56 LQGPAGVGRDSSPGANGIPGTGIPGRDGFGEKRG 91  
DB 1307 DQGPPLQGNPGRPGANGKGDPLGVDGFGPKMG 1342

RESULT 15

T24482  
 hypothetical protein T05A1.2 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
 C:Accession: T24482  
 R;Lloyd, C.  
 submitted to the EMBL Data Library, December 1995  
 A:Reference number: Z19897  
 A:Accession: T24482  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-300 <NLT>  
 A:Cross-references: UNIPROT:Q22183; EMBL:Z68219; PIDN:CAA92476.1; GSPDB:GN00022; CESP:T05A1  
 A:Experimental source: clone T05A1  
 C:Genetics:  
 A:Gene: CESP:T05A1.2  
 A:Map position: 4  
 A:introns: 7/3

Query Match 9.4%; Score 122.5; DB 2; Length 300;  
 Best Local Similarity 31.5%; Pred. No. 0.0021;  
 Matches 35; Conservative 11; Mismatches 28; Indels 37; Gaps 5;  
 QY 3 POGPAASPOR--LRGLLLILLQLPAVSSASEIPKQKQKQLRQREVVDLYNGMCLQGP- 59  
 DB 189 PHGPNGHPGKPKGSQG-----PGPPGHSDPEPKQPGQPGRAGP-----RGPR 231  
 QY 60 --AGVPRGDSGPA-----NGIPRTGIRGRDGFKQKQKGC 93  
 DB 232 GVAGIRGKDQAPGSPGQPGPRGPGEPQDQAPGQPELPESDGTPEKKGIC 282

Search completed: August 4, 2005, 04:22:29  
 Job time : 20.3765 secs

This Page Blank (uspto)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: August 4, 2005, 04:07:24 ; Search time 88.6235 Seconds

(without alignments)  
1404.088 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303

Sequence: 1 MRQGPASQRLRGILL.....GDASTGMSVSRILIELPK 243

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1303	100.0	243	2	096CG8
2	1294	94.3	243	2	06UW91
3	1225	94.0	245	2	08CG08
4	1210	92.9	245	2	09DID6
5	1061	81.4	232	2	08IX63
6	928	71.2	231	2	06AXI0
7	139	10.7	565	2	08K036
8	133.5	10.2	717	2	09N052
9	131.5	10.1	246	1	CIQC.MOUSE
10	131.5	10.1	1472	2	0902A0
11	129	9.9	287	2	08CFR0
12	129	9.9	1752	2	007265
13	128	9.8	289	2	018799
14	127.5	9.8	358	2	06MEY7
15	127.5	9.8	1069	2	06LAN8
16	127.5	9.8	1461	2	076045
17	127.5	9.8	1464	1	CALL_HUMAN
18	127.5	9.8	1464	1	08N473
19	127	9.7	291	2	09NAR3
20	126.5	9.7	540	2	086Y22
21	126.5	9.7	568	2	08CD80
22	126.5	9.7	739	2	070575
23	126.5	9.7	747	2	06NM57
24	126.5	9.7	751	2	09RIN9
25	126	9.7	327	2	001799
26	126	9.7	754	1	CAS4_CANFA
27	126	9.7	1447	2	06PAU1
28	126	9.7	1447	2	06U1U5
29	126	9.7	1659	2	09QZS0
30	125	9.6	551	2	06ZM13
31	125	9.6	886	2	09NUB7

32	125	9.6	888	2	090796	090796	gallus gall
33	125	9.6	1670	1	CAS4_HUMAN	001955	homo sapien
34	125	9.6	1685	1	CAS4_HUMAN	P29400	homo sapien
35	124.5	9.6	289	1	CQT7_HUMAN	Q9BXJ2	homo sapien
36	124.5	9.6	326	2	0677V9	0677V9	lymphocysti
37	124.5	9.6	1336	2	06R241	06R241	homo sapien
38	124.5	9.6	1516	1	CALH_HUMAN	P39060	homo sapien
39	124.5	9.6	1516	2	06R239	06R239	homo sapien
40	124.5	9.6	1684	2	08HYC1	08HYC1	canis fam11
41	124.5	9.6	1691	2	0866Z2	0866Z2	canis fam11
42	124.5	9.6	1751	2	06R240	06R240	homo sapien
43	124	9.5	957	2	096P44	096P44	homo sapien
44	124	9.5	957	2	09H0V3	09H0V3	homo sapien
45	124	9.5	1163	2	08N6U4	08N6U4	homo sapien

## ALIGNMENTS

RESULT 1

ID	Q96CG8	PRELIMINARY;	PRT;	243 AA.
AC	Q96CG8			
DT	01-DEC-2001 (T-EMBLrel. 19, Created)			
DT	01-DEC-2001 (T-EMBLrel. 19, Last sequence update)			
DT	25-OCT-2004 (T-EMBLrel. 28, Last annotation update)			
DE	Collagen triple helix repeat containing 1 (Collagen triple helix repeat-containing protein 1).			
GN	Name=CTHRC1;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Hsieh F., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Muljany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.C., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., Krzyzanski M.I., Skalska U., Smallus D.E., Scherch A., Schein J.E., Jones S.J., Marra M.A.;			
RT	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."			
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RA	Strausberg R.;			
RL	Submitted (SEP-2001) to the EMBL/Genbank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Acorta smooth muscle;			
RA	Lehner W., Moore D.P., Harmon K.J., Mancini M.L., Lindner V.;			
RL	Submitted (JUL-2002) to the EMBL/Genbank/DBJ databases.			
DR	EMBL: BC014245; AA014245.1; -			
DR	EMBL: AY136825; AA015749.1; -			
DR	Gene, HGNC:18631; CTRC1.			
DR	GO: GO:0005737; C:cytoplasm; IEA.			
DR	GO: GO:0006817; P:phosphate transport; IEA.			
DR	InterPro: IPR008161; Clg_helix.			

DR InterPro: IPR008160; Collagen.  
 DR Pfam: PF01391; Collagen; 1.  
 DR ProDom: PD000007; C1g\_helix; 1.  
 KW Collagen.  
 SQ SEQUENCE 243 AA; 26224 MW; 111FEEB1C66867P9 CRC64;

Query Match 100.0%; Score 1303; DB 2; Length 243;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-107;  
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPASPORLRGLLILLLQLPAPSSASEIPKQKQAKLRQREVVLDYNGMCLQGPA 60  
 DB 1 MRPQGPASPORLRGLLILLLQLPAPSSASEIPKQKQAKLRQREVVLDYNGMCLQGPA 60  
 QY 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 DB 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 QY 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 DB 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYL 180  
 DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYL 180  
 QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSR11IEE 240  
 DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSR11IEE 240  
 QY 241 LPK 243  
 DB 241 LPK 243

RESULT 2  
 ID Q6UW91 PRELIMINARY; PRT; 243 AA.  
 AC Q6UW91;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE CTHRC1.  
 GN ORFNames=UNO762;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;  
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,  
 Chen J., Chow B., Chui C., Crowley C., Currell B., Denel B., Dowd P.,  
 Eaton D., Foster J., Grimaldi C., Gu Q., Hase P.E., Heidens S.,  
 Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,  
 Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schenfeld J.,  
 Seshagiri S., Simons L., Singh J., Smith V., Stinson J., Vagts A.,  
 Vanden R., Watanabe C., Wleand D., Woods K., Xie M.H., Yansura D.,  
 Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,  
 RA Godowski P.;  
 RT "The secreted protein discovery initiative (SPDI), a large-scale  
 effort to identify novel human secreted and transmembrane proteins: a  
 bioinformatics assessment";  
 RT bioinformatics assessment";  
 RL Genome Res. 13:2265-2270(2003).  
 DR EMBL: AY358914; AA089273.1; -;  
 DR GO: GO:0005737; C:cytoplasm; IEA.  
 DR GO: GO:0006817; P:phosphate transport; IEA.  
 DR InterPro: IPR008161; C1g\_helix.  
 DR InterPro: IPR008161; C1g\_helix.  
 DR ProDom: PD000007; C1g\_helix; 1.  
 KW Collagen.  
 SQ SEQUENCE 243 AA; 26266 MW; BCB49AF4DBC303BC CRC64;

Query Match 99.3%; Score 1294; DB 2; Length 243;  
 Best Local Similarity 99.6%; Pred. No. 5.7e-106;  
 Matches 242; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRPQGPASPORLRGLLILLLQLPAPSSASEIPKQKQAKLRQREVVLDYNGMCLQGPA 60  
 DB 1 MRPQGPASPORLRGLLILLLQLPAPSSASEIPKQKQAKLRQREVVLDYNGMCLQGPA 60  
 QY 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 DB 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 QY 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 DB 61 GVPGRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 QY 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYL 180  
 DB 121 GKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYL 180  
 QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSR11IEE 240  
 DB 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSR11IEE 240  
 QY 241 LPK 243  
 DB 241 LPK 243

RESULT 3  
 ID Q8CG08 PRELIMINARY; PRT; 245 AA.  
 AC Q8CG08;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Collagen triple helix repeat-containing protein 1.  
 GN Name=Cthrc1;  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Sprague-Dawley; TISSUE=8 day balloon-injured carotid artery;  
 RL Lehnert W., Moore D.P., Harmon K.J., Mancini M.L., Lindner V.,  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: A136824; AAN15746.1; -;  
 DR GO: GO:0005737; C:cytoplasm; IEA.  
 DR GO: GO:0006817; P:phosphate transport; IEA.  
 DR InterPro: IPR008161; C1g\_helix.  
 DR InterPro: IPR008161; C1g\_helix.  
 DR Pfam: PF01391; Collagen.  
 DR ProDom: PD000007; C1g\_helix; 1.  
 KW Collagen.  
 SQ SEQUENCE 245 AA; 26424 MW; 2296FD6DCDBA21F2 CRC64;

Query Match 94.0%; Score 1225; DB 2; Length 245;  
 Best Local Similarity 94.7%; Pred. No. 6.9e-100;  
 Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MRPQGPASPORLRG-LILLLILLLQLPAPSSASEIPKQKQAKLRQREVVLDYNGMCLQ 58  
 DB 1 MRPQGPASPORLRG-LILLLILLLQLPAPSSASEIPKQKQAKLRQREVVLDYNGMCLQ 60  
 QY 59 PAVGVRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 118  
 DB 59 PAVGVRDGSFGANGIPGTGPIGRDGFKEGKECELRSEFSESWTPYKQCSWSSLNVI 120  
 QY 119 DGKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYL 178  
 DB 121 DGKIAECTFTKMSNSALRVLFSGSLRLKCRNACCQRMWTFENGACSGPLPIEAIYYL 180  
 QY 179 DQSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSR11IEE 238  
 DB 181 DQSPENNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMNSVSR11IEE 240  
 QY 239 EELPK 243  
 DB 241 EELPK 245

RESULT 4			
ID	Q9D1D6	PRELIMINARY;	PRT; 245 AA.
AC	Q9D1D6		
DT	01-JUN-2001	(TrEMBLrel. 17, Created)	
DT	01-JUN-2001	(TrEMBLrel. 17, Last sequence update)	
DT	01-JUN-2003	(TrEMBLrel. 24, Last annotation update)	
DE	Mus musculus 16-day embryo whole body cDNA, RIKEN full-length enriched library, clone:111001B07 product:hypothetical Collagen triple helix repeat containing protein, full insert sequence.		
GN	Name=Ctrcl;		
OS	Mus musculus (Mouse).		
OC	Eumetazoa; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Whole body;		
RC	MEDLINE=99279253; PubMed=103449636; DOI=10.1016/S0076-6879(99)03004-9;		
RX	MEDLINE=1085660; PubMed=11217851; DOI=10.1038/35055500;		
RA	RIKEN FANTOM Consortium;		
RT	"Functional annotation of a full-length mouse cDNA collection.";		
RL	Nature 409:685-690(2001).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Whole body;		
RA	The RIKEN Genome Consortium.		
RT	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";		
RL	Nature 420:563-573(2002).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Whole body;		
RX	MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;		
RA	Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,		
RT	"Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";		
RL	Genome Res. 10:1617-1630(2000).		
RN	[5]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Whole body;		
RX	MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;		
RA	Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,		
RA	Kono H., Akiyama J., Nishi K., Kitsuunai T., Tashiro H., Itoh M.,		
RA	Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,		
RA	Yamanoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kaibinagi K.,		
RA	Fujiwaka S., Inoue K., Togawa Y., Iizawa M., Ohara E., Watanabe M.,		
RA	Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunura S., Kawai J.,		
RA	Ozekaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;		
RT	"RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";		
RL	Genome Res. 10:11757-11771(2000).		
RN	[6]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=Whole body;		
RC	Adachi J., Aizawa K., Akahita S., Akimura T., Arai A., Aono H.,		
RA	Akawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,		
RA	Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,		
RA	Imotani K., Ishii Y., Itoh M., Iizawa M., Kasukawa T., Kato H.,		
RA	Kawai Y., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,		
RA	Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,		
RA	Ozekaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,		
RA	Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shitaki T.,		
RA	Sogade Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,		

QY	DB	Query Match	Best Local Similarity	Matches	Conservative	Score	DB 2	Length	DB 2	Indels	Gaps
QY	1	MRPGGPAASQRLRG--LILLLLILQLPAPSSASSTPRKQKQKQRLRQREVDVLLNYGMCILQG	92.9%;	229;	4;	1210;	DB 2;	245;	1;		
DB	1	MHPGGRAPLPQLLGLFLVLLLLQLQALPAPISASNPVKYKQALRLRQREVDVLLNYGMCILQG	93.5%;	229;	4;	1210;	DB 2;	245;	1;		
QY	59	PAGVPGDGSPPGANGIGTGTGPIGPRDGPKEKGECLRESPEESMTPNYKQCSWSLNGI	92.9%;	229;	4;	1210;	DB 2;	245;	1;		
DB	61	PAGVPGDGSPPGANGIGTGTGPIGPRDGPKEKGECLRESPEESMTPNYKQCSWSLNGI	93.5%;	229;	4;	1210;	DB 2;	245;	1;		
QY	119	DLGVIACCTTKMNSNGALRYLFFSGSLRLKCRNACCGRWYTFPFGAECGSLPIEAIYYL	92.9%;	229;	4;	1210;	DB 2;	245;	1;		
DB	121	DLGVIACCTTKMNSNGALRYLFFSGSLRLKCRNACCGRWYTFPFGAECGSLPIEAIYYL	93.5%;	229;	4;	1210;	DB 2;	245;	1;		
QY	179	DQSGPENNSTINIHRTSSVEGLCEGIGAGLVDAIWWGTCSDDYPKGDASTGMSVSRIL	92.9%;	229;	4;	1210;	DB 2;	245;	1;		
DB	181	DQSGPENNSTINIHRTSSVEGLCEGIGAGLVDAIWWGTCSDDYPKGDASTGMSVSRIL	93.5%;	229;	4;	1210;	DB 2;	245;	1;		
QY	239	ELLPK 243	92.9%;	229;	4;	1210;	DB 2;	245;	1;		
DB	241	ELLPK 245	93.5%;	229;	4;	1210;	DB 2;	245;	1;		

	RESULT	5			
OY	ID	Q8IX63	PRELIMINARY;	PRT:	232 AA.
AC	Q8IX63				
DT	01-MAR-2003	(TREMBLrel. 23,	Created)		
DT	01-MAR-2003	(TREMBLrel. 23,	Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25,	Last annotation update)		
DE	NTRYC1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RL	Sanuki N., Fujiki K., Kanai A., Tanaka Y., Iwata T.;				
PL	Submitted (JUN-2001) to the EMBL/GenBank/DDB databases.				
DR	EMBL; AF295488; AA017919.1; -.				
DR	GO; GO:0005737; C:cytosolasm; IEA.				
DR	GO; GO:0006817; P:phosphate transport; IEA.				
DR	InterPro; IPRO08161; Clg_helix.				
DR	InterPro; IPRO08160; Collagen.				
DR	Pfam; PF01391; Collagen; 1.				
DR	Prodom; PD000007; Clg_helix; 1.				
KW	Collagen.				
SQ	SEQUENCE	232 AA;	25163 MW;	E9D4BC30304837ED CRC64;	
	Query Match	81.4%;	Score 1061;	DB 2;	Length 232;
	Best Local Similarity	83.5%;	Pred. No. 1.9e-65;		
	Matches 202;	Conservative 8;	Mismatches 18;	Indels 14;	Gaps 2;
OY	1	MRRQGPAASQRIRGLILLILQLPAPSSASSLPKGGKKRQLAFREVVVDLYNGCLCGPA	60		
db	1	MPPGGRSIT-VKIREKTVSRKLEMGNSASRGILICG-----YNNGCLCGPA	46		

```
QY 61 GVPGRDGPANGIPGTGGIPIGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDL 120
DB 47 GVPGRGSGPANGIPGTGGIPIGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDL 106
QY 121 GKIAECTFTKMRNSALRYLFGSLRLKCRNACCQWYFTFNAGECGSLPIEAIYYLDQ 180
DB 107 GKIAECTFTKMRNSALRYLFGSLRLKCRNACCQWYFTFNAGECGSLPIEAIYYLDQ 166
QY 181 GSEPMNSTINIRHTSVSEGLCEGIGAGLVDAIWMVGTCSDPYKGDASTGMSVSRITIEE 240
DB 167 GSEPMNSTINIRHTSVSEGLCEGIGAGLVDAIWMVGTCSDPYKGDASTGMSVSRITIEE 226
QY 241 LP 242
DB 227 LP 228

RESULT 6
Q6AXL0 PRELIMINARY; PRT; 231 AA.
ID 06AXL0;
AC 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DE 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Zgc:101075 protein.
GN Name=zgc:101075;
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP STRAIN=Singapore local strain; TISSUE=Embryo;
RX PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullanb S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Singapore local strain; TISSUE=Embryo;
RA Director MGC Project;
RA Submitted (Aug-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC079494; AAH79494.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR InterPro; IPR008167; P:phosphate transport; IEA.
DR DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 1.
KW Collagen.
SQ SEQUENCE 231 AA; 25395 MW; 4F9C698080FE1F9B CRC64;
```

```
DB 4 KLTQGLICFWISLPFCVYQKAKERIPR-----QDAFTTKYQA-CVQGVVGVGRGN 56
QY 69 PGANGIPGTGGIPIGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDLKIAECTF 128
DB 57 PGANGIPGTGGIPIGRDGFKEGKECELRSPFESWTPNYKQCSWSLNYGIDLKIAECTF 116
QY 129 TKGRNSALRYLFGSLRLKCRNACCQWYFTFNAGECGSLPIEAIYYLDQSPENNST 188
DB 117 TKGRNSALRYLFGSLRLKCRNACCQWYFTFNAGECGSLPIEAIYYLDQSPENNST 176
QY 189 INIRHTSVSEGLCEGIGAGLVDAIWMVGTCSDPYKGDASTGMSVSRITIEE 243
DB 177 INIRHTSVSEGLCEGIGAGLVDAIWMVGTCSDPYKGDASTGMSVSRITIEE 231

RESULT 7
Q8K036 PRELIMINARY; PRT; 565 AA.
ID 08K036;
AC 08K036;
DT 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DE 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Col13a1 protein.
GN Name=Col13a1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullanb S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;
RA Strausberg R.;
RA Submitted (Jul-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC034164; AAH34164.1; -.
DR MGD; MGI:1277201; Col13a1.
DR GO; GO:0005911; C:intercellular junction; IDA.
DR InterPro; IPR008161; C1g_helix.
DR DR InterPro; IPR008160; Collagen.
DR Pfam; PF01391; Collagen; 5.
DR ProDom; PD000007; C1g_helix; 1.
KW Collagen.
SQ SEQUENCE 565 AA; 56726 MW; DBD3F99D670195F CRC64;
```

Query Match 71.2%; Score 928; DB 2; Length 231;  
Best Local Similarity 72.8%; Pred. No. 1e-73;  
Matches 111; Conservative 21; Mismatches 33; Indels 10; Gaps 3;

Query Match 10.7%; Score 139; DB 2; Length 565;  
Best Local Similarity 34.8%; Pred. No. 0.0011;  
Matches 39; Conservative 12; Mismatches 35; Indels 26; Gaps 6;



```

Db      393  PGAAGGCGSPGPG-AKGEPRKGMVD-YNGSINELAOEIRITLALMPPELPQGTGPRG 450
Qy      72  NGIP-----GTTPGPRDGFKEBKGECLRESEESTPNYKCCSSWLVNG 117
Db      451  PGTGPRGGEIGLPGPPGHDGKPRG---KPGDAGWSSTPRKGMASRNEG 499

RESULT 8
Q9N052  PRELIMINARY; PRT; 717 AA.
ID      Q9N052
AC      Q9N052;
DT      01-OCT-2000 (T-EMBLrel. 15, Created)
DT      01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT      01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE      Type XIII collagen.
GN      Name=COLXIII.1;
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Snellman A., Tu H., Vaeisaeen T., Kvist A.P., Huhtala P.,
RA      Phlajantemi T.;
RL      Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ293624; CAC00688.1; -.
DR      Genew; HGNC:12190; COL13A1.
DR      GO; GO:0005737; Cytoplasm; IEA.
DR      GO; GO:0006817; P-transport; IEA.
DR      InterPro; IPR008161; C1q helix.
DR      InterPro; IPR008160; Collagen.
DR      Pfam; PF01391; Collagen_9.
DR      ProDom; PD000007; C1q_helix; 3.
KW      Collagen.
SQ      SEQUENCE 717 AA; 69964 MW; A311B9C7D3E87577 CRC64;

Query Match 10.2%; Score 133.5; DB 2; Length 717;
Best Local Similarity 43.4%; Pred. No. 0.0045;
Matches 33; Conservative 8; Mismatches 16; Indels 19; Gaps 4;

Qy      34  PKGKQKALQREVDLYNG-----MCLQGPAGVPGRDSPGANGIP-----G 76
Db      431  PKG-SKGEPRKGMVD-YNGSINELAOEIRITLALMPPELPQGTGPRG 488

Qy      77  TPGIPGRDGFKEBKGE 92
Db      489  LFGPPGHDGKPRGK 504

RESULT 9
C1QC_MOUSE STANDARD; PRT; 246 AA.
ID      C1QC_MOUSE
AC      Q02105;
DT      01-JUL-1993 (Rel. 26, Created)
DT      01-JUL-1993 (Rel. 26, Last sequence update)
DT      25-JAN-2005 (Rel. 46, Last annotation update)
DE      Complement C1q subcomponent, C chain precursor.
GN      Name=C1q; Synonyms=C1q;
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_TaxID=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=ICR; TISSUE=Macrophage;
RX      MEDLINE=9301118; PubMed=1396691;
RA      Petry F., Reid K.B.M., Loos M.;
RA      Petry F., Reid K.B.M., Loos M.;
RT      "Isolation, sequence analysis and characterization of cDNA clones
RT      coding for the C chain of mouse C1q. Sequence similarity of complement
RT      subcomponent C1q, collagen type VIII and type X and precerebellin.",
RL      Eur. J. Biochem. 209:129-134(1992).
RN      [2]

```

```

RP      SEQUENCE FROM N.A.
RC      STRAIN=BALB/c; TISSUE=Liver;
RX      MEDLINE=96186528; PubMed=8606057; DOI=10.1007/s002510050077;
RA      Petry F., McClive P.J., Bolto M., Morley B.J., Moran G., Loos M.;
RT      "The mouse C1q genes are clustered on chromosome 4 and show
RT      conservation of gene organization.";
RL      Immunogenetics 43:370-376(1996).
RN      [3]
RN      SEQUENCE FROM N.A.
RP      STRAIN=FVB/N; TISSUE=Colon;
RX      MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA      Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA      Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schler G.D.,
RA      Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA      Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA      Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA      Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
RA      Brownstein M.J., Umed T.B., Toshiyuki S., Carninci P., Prange C.,
RA      Raha S.S., Loughran N.A., Peters G.J., Abramson R.D., Millar S.J.,
RA      Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA      Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA      Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA      Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA      Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA      Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA      Butcherfield Y.S.N., Krzywicki M.I., Skalska U., Smalins D.E.,
RA      Scherkl J.E., Schein J.E., Jones S.J.M., Marra M.A.;
RT      "Generation and initial analysis of more than 15,000 full-length human
RT      and mouse cDNA sequences.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN      [4]
RN      FUNCTION: C1q associates with the proenzymes C1r and C1s to yield
RN      C1, the first component of the serum complement system. The
RN      collagen-like regions of C1q interact with the Ca(2+)-dependent
RN      C1r(2)C1s(2) proenzyme complex, and efficient activation of C1
RN      takes place on interaction of the globular heads of C1q with the
RN      FC regions of IgG or IgM antibody present in immune complexes.
RN      -1- SUBUNIT: C1 is a calcium-dependent trimolecular complex of C1q, R
RN      and S in the molar ratio of 1:2:2. C1q subcomponent is composed
RN      of nine subunits, six of which are disulfide-linked dimers of the
RN      A and B chains, and three of which are disulfide-linked dimers of
RN      the C chain.
RN      -1- SIMILARITY: Contains 1 C1q domain.
RN      -1- SIMILARITY: Contains 1 collagenous domain.
RN      This SWISS-PROT entry is copyright. It is produced through a collaboration
RN      between the Swiss Institute of Bioinformatics and the EMBL outstation -
RN      the European Bioinformatics Institute. There are no restrictions on its
RN      use by non-profit institutions as long as its content is in no way
RN      modified and this statement is not removed. Usage by and for commercial
RN      entities requires a license agreement (See http://www.isb-sib.ch/announce/
RN      or send an email to license@sib-sib.ch).
RN      -----
RN      EMBL; X66295; CAA46993.1; -.
RN      EMBL; X92960; CAA63535.1; -.
RN      EMBL; BC054443; AAH54443.1; -.
RN      F1R; S29328; S29328.
RN      HSSP; G60994; IC28.
RN      MGD; MG1:88225; C1qg.
RN      InterPro; IPR001073; C1q.
RN      InterPro; IPR008161; C1q helix.
RN      InterPro; IPR008160; Collagen.
RN      InterPro; IPR008963; TNF-like.
RN      Pfam; PF00386; C1q; 1.
RN      Pfam; PF01391; Collagen; 1.
RN      PRINTS; PR00007; COMPLEMENTC1Q.
RN      ProDom; PD000007; C1q helix; 1.
RN      PROSITE; PS50871; C1Q_1.
RN      Collagen; Complement pathway; Glycoprotein; Hydroxylation; Plasma;
RN      Repeat; Signal.
RN      SIGNAL 1 29
RN      CHAIN 30 246 Complement C1q subcomponent, C chain.
RN      FT 32 113 Collagen-like.
RN      DOMAIN 32 113

```

```
FT DOMAIN 116 246 Ctg.
FT DISULFID 33 33 Interchain (with other C chain) (By
FT MOD RES 37 37 similarity):
FT MOD RES 40 40 Hydroxyproline (By similarity).
FT MOD RES 43 43 Hydroxyproline (By similarity).
FT MOD RES 61 61 Hydroxyproline (By similarity).
FT MOD RES 64 64 Hydroxyproline (By similarity).
FT MOD RES 72 72 5-Hydroxylysine (By similarity).
FT MOD RES 76 76 5-Hydroxylysine (By similarity).
FT MOD RES 79 79 Hydroxyproline (By similarity).
FT MOD RES 82 82 Hydroxyproline (By similarity).
FT MOD RES 97 97 Hydroxyproline (By similarity).
FT MOD RES 100 100 Hydroxyproline (By similarity).
FT MOD RES 106 106 Hydroxyproline (By similarity).
FT MOD RES 109 109 Hydroxyproline (By similarity).
SQ SEQUENCE 246 AA; 25966 MW; 2F79EA12748CB8E0 CRC64;

Query Match 10.1%; Score 131.5; DB 1; Length 246;
Best Local Similarity 38.5%; Pred. No. 0.0019;
Matches 35; Conservative 5; Mismatches 32; Indels 19; Gaps 2;

QY 5 GPAPSPQRRLGILLILLQLPAPSSASEIPKGRKQAQLRQREVVDLYNQMCIQGPAGVGG 64
DB 4 GPSCQPGCGICLLILFLILALPLRSQS-----AGCYGIPGMGMGAGG 47
QY 65 RDGS---PGANGIPGTGIPGRDGFKEGKE 92
DB 48 KDGHDGLQGPKEPGIPAVPGTQPGKQGE 78

RESULT 10
Q90ZAO PRELIMINARY; PRT; 1472 AA.
AC Q90ZAO;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Collagen type XX alpha 1 precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21303548; PubMed=11274142; DOI=10.1074/jbc.M009912200;
RA Koch M., Foley J.E., Hahn R., Burgess R.E., Gerecke D.R.,
RA Gordon M.K.;
RT "alpha 1(X) collagen, a new member of the collagen subfamily, fibril-
RT associated collagens with interrupted triple helices."
RL J. Biol. Chem. 276:23120-23126(2001).
DR EMBL; AF312825; AAK58847.1; -.
DR HSSP; Q96KP7; 1FNA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR Pfam; PF01391; Collagen; 4.
DR Pfam; PF00041; FN3; 6.
DR Pfam; PF00092; VWA; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR ProDom; PD000007; C1g_helix; 4.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00210; TSEN; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50853; FN3; 6.
DR PROSITE; PS50234; VWFA; 1.
KW Collagen; Signal.
FT SIGNAL 1 28 Potential.
SQ SEQUENCE 1472 AA; 156903 MW; 5361611579C56EFD CRC64;

Query Match 10.1%; Score 131.5; DB 2; Length 1472;
```

```
Best Local Similarity 34.5%; Pred. No. 0.016;
Matches 38; Conservative 7; Mismatches 26; Indels 39; Gaps 6;

QY 3 PGAPSPQRRLGILLILLQLP-----APSSASEI---PGKQAQLRQREV 47
DB 1161 PGPGSPGGR-RG-----PGGEGEPGPKPGPPGKGPAGPGCGQGSQSG--- 1207
QY 48 VDLNVMGLQGPAGVP-----GRDSPGANGIPGTGIPGRDGFKEGKE 91
DB 1208 -----GITTIQGVPPGIKEKGTGSPMGGITPVGQAPGAGDGLQAGK 1252

RESULT 11
Q9CER0 PRELIMINARY; PRT; 287 AA.
AC Q9CER0;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Similar to C1g-like.
GN Name=BC040774;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carrinck P., Prange C.,
RA Raha S.S., Loughell N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
RA Villalon D.K., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Buterfield Y.S.,
RA Krzywinski M.T., Skalska U., Smalins D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RT Submitted (Nov-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; BC040774; AAH40774.1; -.
DR HSSP; O60994; 1C28.
DR MGD; MGI:3032521; BC040774.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR001073; C1g.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; C1g; 1.
DR Pfam; PF01391; Collagen; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
KW Collagen.
SQ SEQUENCE 287 AA; 29292 MW; 8FF89BC1C7420415 CRC64;

Query Match 9.9%; Score 129; DB 2; Length 287;
Best Local Similarity 35.5%; Pred. No. 0.0039;
Matches 43; Conservative 10; Mismatches 56; Indels 12; Gaps 4;
```

```

QY 27 PSSASEIPKQKQAQLRQREV-DLYNG---MCIQGPAGVPGRDGSGANGIRGTGCTGPG 82
DB 44 PAKAKAPPPSTAL---EVMODLSANPPPIQGGKGPGRKPGPGPGPGPG 100
QY 83 RDGFKGEGCELRSEFSESWTPYKQCSGLNYGIDLGKIACTFTYKMSNGLRVLPFS 142
DB 101 PRPPPEKGDGSGRPGU-----PGLQLTSAAGGVVSGGTGGGGDTGSEVTSLNLSAAPS 155
QY 143 G 143
DB 156 G 156

RESULT 12
Q07265 PRELIMINARY; PRT; 1752 AA.
ID 007265;
AC 007265;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE 3 alpha procollagen.
GN Name=COLP3alpha;
OS Strongylocentrotus purpuratus (Purple sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidea; Echinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC Strongylocentrotus.
OX NCBI_TaxID=7668;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93186842; PubMed=8444899;
RA Exposito J.-Y., D'Alessio M., Di Liberto M., Ramirez F.;
RT "Complete primary structure of a sea-urchin type IV collagen and
RT analysis of the 5' end of its gene."
RL J. Biol. Chem. 268:5249-5254(1993).
DR EMBL; L02917; AAA30039.1; -.
DR HSSP; P02462; 1L11.
DR GO; GO:0005581; C:collagen; IEA.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0005201; F:extracellular matrix structural constituent, IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR001442; Procollagen4_C.
DR Pfam; PF01413; C4; 2.
DR Pfam; PF01391; Collagen; 25.
DR SMART; SM00111; C4; 2.
DR COLLAGEN.
KW SMART.
SQ SEQUENCE 1752 AA; 170210 MW; 1AE5AA21569346D CRC64;

Query Match 9.9%; Score 129; DB 2; Length 1752;
Best Local Similarity 22.9%; Pred. No. 0.032;
Matches 47; Conservative 20; Mismatches 60; Indels 78; Gaps 9;

QY 3 PGPASAPORLRLILLILLQLPAPSSASEIPKQKQAQLRQREVVDLYNGMCLQSPAGV 62
DB 1453 PGCPFRDGR-----PGPG-----PG-----LTGDKGT 1477
QY 63 PGHDGSPGANGIPGTGPIGRDGFKEGKECELRSEF-----ESWT 103
DB 1478 PGVQGNPRGVSGVGEGLKGEQFRQNGPGRPGPTKGEAGIRPGSSSGFTTHSGT 1537
QY 104 PNYKQ-----SW-----SLNYGIDLGKIACTFTYKMSNGLRVLPFSGLRL 147
DB 1538 TSPQCPQGTAKMNGVSLFVQGNRGRGQDLGKPGSC-----LKRFTWPLFC-NINN 1592
QY 148 KCRNACCORWYFTFNGACSGPLPI 172
DB 1593 VCHVASRNDYSLWLTSTT---PWPM 1614

RESULT 13
Q18799 PRELIMINARY; PRT; 289 AA.
ID Q18799

```

```

AC Q18799;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C53B4.5.
GN ORFNames=C53B4.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Bristol N2;
RA Bérks M.;
RL Submitted (DEC-1995) to the EMBL/Genbank/DBJ databases.
DR EMBL; Z68215; CAA92453.1; -.
DR PIR; T20177; T20177.
DR WormBase; WBGene0000693; C53B4.5.
DR WormPep; C53B4.5; C603091.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0042302; F:structural constituent of cuticle; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF01484; Col_cuticle_N; 1.
DR COLLAGEN; Hypothetical protein.
KW COLLAGEN.
SQ SEQUENCE 289 AA; 28993 MW; 2999A3FF9CC0B044 CRC64;

Query Match 9.8%; Score 128; DB 2; Length 289;
Best Local Similarity 35.7%; Pred. No. 0.0048;
Matches 30; Conservative 5; Mismatches 19; Indels 30; Gaps 2;

QY 25 PPSASEIPKQKQAQLRQREVVDLYNGMCLQSPAGVGRDGS-----P 69
DB 203 PPGSFGEPGRPGPGS-----RGPAGQKQAGCGPGEKANGEPGP 247
QY 70 GAGIPGTGPIGRDGFKEGKEC 93
DB 248 GRDQGRGRGCGRGRGHEKGV 271

RESULT 14
Q06MEY7 PRELIMINARY; PRT; 358 AA.
ID Q06MEY7;
AC Q06MEY7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN Hypothetical protein.
OS Orderedcushnemes-pc0138;
OC Parachlamydia sp. (strain UME25) (subsp. Acanthamoeba sp.).
OC Bacteria; Chlamydiae; Parachlamydiae; Parachlamydiaceae; Parachlamydia.
OX NCBI_TaxID=264201;
RN [1]
RP SEQUENCE FROM N.A.
RA Horn M., Collingro A., Schmitz-Esser S., Beier C.L., Purkhold U.,
RA Fartmann B., Brandt P., Nyakatura G.J., Droge M., Friehman D.,
RA Rattei T., Mewes H.-W., Wagner M.;
RT "Genome sequence of an amoeba symbiont and its use for reconstructing
RT the evolutionary history of chlamydiae."
RL Submitted (JAN-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; BX908798; CAF22862.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.

```

DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen; 1.  
KW Complete proteome; Hypothetical protein.  
SQ SEQUENCE 358 AA; 37332 MW; 87419BD361E3D618 CRC64;

Query Match 9.8%; Score 127.5; DB 2; Length 358;

Best Local Similarity 28.8%; Pred. No. 0.0068;

Matches 46; Conservative 15; Mismatches 72; Indels 27; Gaps 5;

QY 34 PKGKQAQLRQREVDLYNGMCIQGPAGVPGRDSPGANGIPGTPGIPGRDGFKEKG-- 91  
DB 136 PPGPGQATPGPGPPGPTGAGLTGPTGLTGATGLTGADGFPADGWSGSKGLT 195  
QY 92 -----ECLRSFESWTPNY-----KQCSWSLNYGIDLGKIA-----CTF 128  
DB 196 GATGLTGATGLTGASGPAFGANNVFAVGIETGSVNSPVNLSFDVCPILDGWRPTNTTF 255  
QY 129 TTKRNSALRLVLFSGSLRLKCRNACCQRWYTFNGACSG 168  
DB 256 TCKQTGYL-VQVGGFTLVMDATCALW-ATFNGSQVEG 293

## RESULT 15

Q6LAN8 PRELIMINARY; PRT; 1069 AA.  
AC O6LAN8;  
DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
DE Collagen type I alpha 1 (Fragment).  
GN Name=COL1A1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA O'Brien K.P.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL; X98705; CAA67261.1; -.  
DR GO; GO:0005737; Cytoplasm; IEA.  
DR GO; GO:006817; Phosphate transport; IEA.  
DR InterPro; IPR008161; Clg\_helix.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR009041; PMP\_SGCL.  
DR Pfam; PF01391; Collagen; 16.  
DR Pfam; PF00093; WWC; 1.  
DR ProDom; PD000007; Clg\_helix; 3.  
DR SMART; SM00214; WWC; 1.  
DR PROSITE; PS01208; WWC\_1; 1.  
DR PROSITE; PS50184; WWC\_2; 1.  
KW Collagen.  
FT NON TER 1069  
SQ SEQUENCE 1069 AA; 97445 MW; EE279B10572FB980 CRC64;

Query Match 9.8%; Score 127.5; DB 2; Length 1069;

Best Local Similarity 37.4%; Pred. No. 0.024;

Matches 34; Conservative 2; Mismatches 22; Indels 33; Gaps 4;

QY 3 PGGPAA--PGRKGLLLLLQLPAPSAASEIPKQKQAQLRQREVDLYNGMCIQGPA 60  
DB 357 PGGPAA--PGRKGLLLLLQLPAPSAASEIPKQKQAQLRQREVDLYNGMCIQGPA 385

QY 61 GVGGRDGPANGIPGTPGIPGRDGFKEKG 91  
DB 386 GVGGRDGPANGIPGTPGIPGRDGFKEKG 416

Search completed: August 4, 2005, 04:21:23  
Job time : 89.6235 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:21:44 ; Search time 86.0824 Seconds  
(without alignments)  
1099.816 Million cell updates/sec

Title: US-10-634-108-4

Perfect score: 1303

Sequence: 1 MRPQGPAPSPQRLRGLLL.....GDASTGWSVRIIEELPK 243

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

1: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubppa/US07\_PUBCOMB.pep:\*  
3: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
5: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*  
6: /cgn2\_6/ptodata/2/pubppa/US06\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubppa/US09\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubppa/US10\_PUBCOMB.pep:\*  
19: /cgn2\_6/ptodata/2/pubppa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/2/pubppa/US11\_PUBCOMB.pep:\*  
21: /cgn2\_6/ptodata/2/pubppa/US11\_PUBCOMB.pep:\*  
22: /cgn2\_6/ptodata/2/pubppa/US11\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1303	100.0	243	US-09-829-472A-19	Sequence 19, Appl
2	1303	100.0	243	US-10-045-992-4	Sequence 4, Appl
3	1303	100.0	243	US-10-097-340-135	Sequence 135, Appl
4	1303	100.0	243	US-10-295-027-50	Sequence 50, Appl
5	1303	100.0	243	US-10-295-027-1173	Sequence 1173, Appl
6	1303	100.0	243	US-10-173-999-74	Sequence 74, Appl
7	1303	100.0	243	US-10-058-270A-38	Sequence 38, Appl
8	1303	100.0	243	US-10-634-108-4	Sequence 4, Appl
9	1303	100.0	243	US-10-188-833-175	Sequence 175, Appl
10	1303	100.0	243	US-10-939-233-4	Sequence 4, Appl
11	1303	100.0	278	US-10-097-340-131	Sequence 131, Appl

12	1303	100.0	278	US-10-177-293-496	Sequence 496, Appl
13	1303	100.0	278	US-10-301-832-77	Sequence 77, Appl
14	1303	100.0	278	US-10-296-115-1261	Sequence 1261, Appl
15	1303	100.0	278	US-10-961-139-2	Sequence 2, Appl
16	1300	99.8	243	US-10-351-334-205	Sequence 205, Appl
17	1300	99.8	278	US-10-060-036-4551	Sequence 4551, Appl
18	1298	99.6	243	US-10-176-847-14	Sequence 14, Appl
19	1294	99.3	243	US-09-834-755-514	Sequence 514, Appl
20	1294	99.3	243	US-09-938-418-7	Sequence 7, Appl
21	1294	99.3	243	US-09-946-374-352	Sequence 352, Appl
22	1294	99.3	243	US-09-946-374-431	Sequence 431, Appl
23	1294	99.3	243	US-10-006-867-122	Sequence 122, Appl
24	1294	99.3	243	US-10-052-586-366	Sequence 366, Appl
25	1294	99.3	243	US-10-007-805-514	Sequence 514, Appl
26	1294	99.3	243	US-10-063-587-122	Sequence 122, Appl
27	1294	99.3	243	US-10-063-551-122	Sequence 122, Appl
28	1294	99.3	243	US-10-174-550-366	Sequence 366, Appl
29	1294	99.3	243	US-10-176-758-366	Sequence 366, Appl
30	1294	99.3	243	US-10-175-737-366	Sequence 366, Appl
31	1294	99.3	243	US-10-063-616-122	Sequence 122, Appl
32	1294	99.3	243	US-10-174-581-366	Sequence 366, Appl
33	1294	99.3	243	US-10-176-483-366	Sequence 366, Appl
34	1294	99.3	243	US-10-176-749-366	Sequence 366, Appl
35	1294	99.3	243	US-10-176-914-366	Sequence 366, Appl
36	1294	99.3	243	US-10-176-915-366	Sequence 366, Appl
37	1294	99.3	243	US-10-063-569-122	Sequence 122, Appl
38	1294	99.3	243	US-10-063-513-122	Sequence 122, Appl
39	1294	99.3	243	US-10-063-515-122	Sequence 122, Appl
40	1294	99.3	243	US-10-063-512-122	Sequence 122, Appl
41	1294	99.3	243	US-10-173-706-366	Sequence 366, Appl
42	1294	99.3	243	US-10-175-728-366	Sequence 366, Appl
43	1294	99.3	243	US-10-175-752-366	Sequence 366, Appl
44	1294	99.3	243	US-10-176-482-366	Sequence 366, Appl
45	1294	99.3	243	US-10-176-757-366	Sequence 366, Appl

#### ALIGNMENTS

RESULT 1  
US-09-829-472A-19  
; Sequence 19, Application US/09829472A  
; Publication No. US20040146862A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David  
; APPLICANT: Gish, Kurt  
; APPLICANT: EOS Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and  
; TITLE OF INVENTION: Methods of Screening for Modulators of Breast Cancer  
; FILE REFERENCE: 018501-001200US  
; CURRENT APPLICATION NUMBER: US/09/829,472A  
; PRIOR FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 09/525,361  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: BCM4, ESTs, secreted protein  
US-09-829-472A-19

Query Match 100.0%; Score 1303; DB 12; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1.56-115;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPQGPAPSPQRLRGLLLIOLPAPSSASIRPKGKAKQRLRQREVDLYNGKLOGPA 60  
DB 1 MRPQGPAPSPQRLRGLLLIOLPAPSSASIRPKGKAKQRLRQREVDLYNGKLOGPA 60  
QY 61 GVPGRDGSFGANGIRGTGIRPORDFKGEKGECLRESPEESWTPTYKQCSNLSNYGIDL 120

```
Db      61 |||||
GVPGRGSGPGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPTYKQCSWSLNYGIDL 120
QY      121 GKIACCTFTKMSNSALRYLFSGSLRLKCRNACCORWYTFNGAECGPLEIATIIYDQ 180
      121 GKIACCTFTKMSNSALRYLFSGSLRLKCRNACCORWYTFNGAECGPLEIATIIYDQ 180
Db      121 GKIACCTFTKMSNSALRYLFSGSLRLKCRNACCORWYTFNGAECGPLEIATIIYDQ 180
QY      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDYPKGDASTGMNSVSRITIEE 240
      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDYPKGDASTGMNSVSRITIEE 240
Db      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDYPKGDASTGMNSVSRITIEE 240
QY      241 LPK 243
      241 LPK 243
Db      241 LPK 243

RESULT 2
US-10-045-992-4
; Sequence 4, Application US/10045992
; Publication No. US20020161211A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN
; FILE REFERENCE: 053689-5006-01
; CURRENT APPLICATION NUMBER: US/10/045,992
; PRIOR FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: 09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-045-992-4

Query Match      100.0%; Score 1303; DB 13; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPOGPAASPORLRGLLILLQLPAPSSASEIPKQKQAKQLRQREVVDLYNGMCLQCPA 60
Db      1 MRPOGPAASPORLRGLLILLQLPAPSSASEIPKQKQAKQLRQREVVDLYNGMCLQCPA 60
QY      61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPTYKQCSWSLNYGIDL 120
Db      61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPTYKQCSWSLNYGIDL 120
QY      121 GKIACCTFTKMSNSALRYLFSGSLRLKCRNACCORWYTFNGAECGPLEIATIIYDQ 180
Db      121 GKIACCTFTKMSNSALRYLFSGSLRLKCRNACCORWYTFNGAECGPLEIATIIYDQ 180
QY      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDYPKGDASTGMNSVSRITIEE 240
Db      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDYPKGDASTGMNSVSRITIEE 240
QY      241 LPK 243
      241 LPK 243
Db      241 LPK 243

RESULT 3
US-10-097-340-135
; Sequence 135, Application US/10097340
; Publication No. US20030087250A1
; GENERAL INFORMATION:
; APPLICANT: John MONAHAN
; APPLICANT: Manjula GANNAVARAPU
; APPLICANT: Sebastian HOERSCHE
; APPLICANT: Subhangt KAMATYAR
; APPLICANT: Steve G. KOVATS
; APPLICANT: Rachel E. MEYERS
```

```
; APPLICANT: Michael MORRISEY
; APPLICANT: Peter OLANDT
; APPLICANT: Ami SEN
; APPLICANT: Peter VEIBY
; APPLICANT: Gordon B. MILLS
; APPLICANT: Robert C. BAST, Jr.
; APPLICANT: Karen LU
; APPLICANT: Rosemarie SCHMANT
; APPLICANT: Xumei ZHAO
; APPLICANT: Karen GLATT
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,
; FILE REFERENCE: MRI-030
; CURRENT APPLICATION NUMBER: US/10/097,340
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 60/276,025
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/325,149
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/276,026
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 60/324,967
; PRIOR FILING DATE: 2001/09/26
; PRIOR APPLICATION NUMBER: 60/311,732
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: 60/325,102
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 60/323,580
; PRIOR FILING DATE: 2001-09-19
; NUMBER OF SEQ ID NOS: 363
; SOFTWARE: FaestSeq for Windows Version 4.0
; SEQ ID NO 135
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-097-340-135

Query Match      100.0%; Score 1303; DB 14; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPOGPAASPORLRGLLILLQLPAPSSASEIPKQKQAKQLRQREVVDLYNGMCLQCPA 60
Db      1 MRPOGPAASPORLRGLLILLQLPAPSSASEIPKQKQAKQLRQREVVDLYNGMCLQCPA 60
QY      61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPTYKQCSWSLNYGIDL 120
Db      61 GVPGRDGSFGANGIPETPGIPGRDGFKEGKEGECLERESFEESWTPTYKQCSWSLNYGIDL 120
QY      121 GKIACCTFTKMSNSALRYLFSGSLRLKCRNACCORWYTFNGAECGPLEIATIIYDQ 180
Db      121 GKIACCTFTKMSNSALRYLFSGSLRLKCRNACCORWYTFNGAECGPLEIATIIYDQ 180
QY      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDYPKGDASTGMNSVSRITIEE 240
Db      181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIHWGTCSDYPKGDASTGMNSVSRITIEE 240
QY      241 LPK 243
      241 LPK 243
Db      241 LPK 243

RESULT 4
US-10-295-027-50
; Sequence 50, Application US/10295027
; Publication No. US2003023250A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natacha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
```



```

; SOFTWARE:Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-270A-38

Query Match          100.0%; Score 1303; DB 15; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 MRPGPASPRLRGLLLLLQLPAPSSASRIPIKQKQAOIRQREVVLDLYNGMCIQGPA 60
DB      1 MRPGPASPRLRGLLLLLQLPAPSSASRIPIKQKQAOIRQREVVLDLYNGMCIQGPA 60

QY      61 GVPGRDGSFGANGIPGTGPIPGRDGFKGEKGECLRESFEESWTPTYKQCSWSLNYGIDL 120
DB      61 GVPGRDGSFGANGIPGTGPIPGRDGFKGEKGECLRESFEESWTPTYKQCSWSLNYGIDL 120

QY      121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFENGACSGPLPIEAIITYLDQ 180
DB      121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFENGACSGPLPIEAIITYLDQ 180

QY      181 GSPEMNSTINIHRTSSVVEGLCEGIGAGLVDAIWWGTGSDPYKGAStGMSNSVSRRIIE 240
DB      181 GSPEMNSTINIHRTSSVVEGLCEGIGAGLVDAIWWGTGSDPYKGAStGMSNSVSRRIIE 240

RESULT 8
US-10-634-108-4
; Sequence 4, Application US/10634108
; Publication No. US20040063176A1
; GENERAL INFORMATION:
; APPLICANT: LINDER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Linder et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/10/634.108
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/692.081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-108-4

Query Match          100.0%; Score 1303; DB 15; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.5e-115;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY      1 MRPGPASPRLRGLLLLLQLPAPSSASRIPIKQKQAOIRQREVVLDLYNGMCIQGPA 60
DB      1 MRPGPASPRLRGLLLLLQLPAPSSASRIPIKQKQAOIRQREVVLDLYNGMCIQGPA 60

QY      61 GVPGRDGSFGANGIPGTGPIPGRDGFKGEKGECLRESFEESWTPTYKQCSWSLNYGIDL 120
DB      61 GVPGRDGSFGANGIPGTGPIPGRDGFKGEKGECLRESFEESWTPTYKQCSWSLNYGIDL 120

QY      121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFENGACSGPLPIEAIITYLDQ 180
DB      121 GKIAECTTKMRSNSALRVLFSGSLRLKCRNACCORWYFTFENGACSGPLPIEAIITYLDQ 180

QY      181 GSPEMNSTINIHRTSSVVEGLCEGIGAGLVDAIWWGTGSDPYKGAStGMSNSVSRRIIE 240
DB      181 GSPEMNSTINIHRTSSVVEGLCEGIGAGLVDAIWWGTGSDPYKGAStGMSNSVSRRIIE 240

```



QY 241 LPK 243  
Db 241 LPK 243

## RESULT 9

US-10-188-832-175  
; Sequence 175, Application US/10188832  
; Publication No. US20040076955A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David H.  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Eos Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions  
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder  
; TITLE OF INVENTION: Cancer  
; FILE REFERENCE: 018501-002330US  
; CURRENT APPLICATION NUMBER: US/10/188,832  
; CURRENT FILING DATE: 2002-11-22  
; PRIOR APPLICATION NUMBER: US 60/302,814  
; PRIOR FILING DATE: 2001-07-03  
; PRIOR APPLICATION NUMBER: US 60/310,099  
; PRIOR FILING DATE: 2001-08-03  
; PRIOR APPLICATION NUMBER: US 60/343,705  
; PRIOR FILING DATE: 2001-11-08  
; PRIOR APPLICATION NUMBER: US 60/350,666  
; PRIOR FILING DATE: 2001-11-13  
; PRIOR APPLICATION NUMBER: US 60/372,246  
; PRIOR FILING DATE: 2002-04-12  
; NUMBER OF SEQ ID NOS: 207  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 175  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-832-175

Query Match 100.0%; Score 1303; DB 15; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1.5e-115;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQLRQREVVDLYNMGCLQGPA 60  
Db 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQLRQREVVDLYNMGCLQGPA 60  
QY 61 GVPGRDGPANGIPGTGGIPGRDGFKEGKEGECLEAESFEESWTPNTYKQCSWSSLNYGIDL 120  
Db 61 GVPGRDGPANGIPGTGGIPGRDGFKEGKEGECLEAESFEESWTPNTYKQCSWSSLNYGIDL 120  
QY 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
Db 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
QY 181 GSPENNSTINIHRTSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNSVSRITIEE 240  
Db 181 GSPENNSTINIHRTSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNSVSRITIEE 240  
QY 241 LPK 243  
Db 241 LPK 243

## RESULT 10

US-10-939-233-4  
; Sequence 4, Application US/10939233  
; Publication No. US20050147602A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volhard  
; APPLICANT: FRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTRHCL, A NOVEL  
; FILE REFERENCE: 053689-5006-03  
; CURRENT APPLICATION NUMBER: US/10/939,233

; CURRENT FILING DATE: 2004-09-10  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US 10/045,992  
; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 4  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-939-233-4

Query Match 100.0%; Score 1303; DB 16; Length 243;  
Best Local Similarity 100.0%; Pred. No. 1.5e-115;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQLRQREVVDLYNMGCLQGPA 60  
Db 1 MRPGPAASPORLRLGLLLLLQLPAPSSASEIPKQKQAKQLRQREVVDLYNMGCLQGPA 60  
QY 61 GVPGRDGPANGIPGTGGIPGRDGFKEGKEGECLEAESFEESWTPNTYKQCSWSSLNYGIDL 120  
Db 61 GVPGRDGPANGIPGTGGIPGRDGFKEGKEGECLEAESFEESWTPNTYKQCSWSSLNYGIDL 120  
QY 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
Db 121 GKIAECTFTKMSNSALRYLFGSGLRLKCRNACCCQWYTFNGAECSSGPLPIAIIYLDQ 180  
QY 181 GSPENNSTINIHRTSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNSVSRITIEE 240  
Db 181 GSPENNSTINIHRTSVEGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNSVSRITIEE 240  
QY 241 LPK 243  
Db 241 LPK 243

## RESULT 11

US-10-097-340-131  
; Sequence 131, Application US/10097340  
; Publication No. US20030087250A1  
; GENERAL INFORMATION:  
; APPLICANT: John MONAHAN  
; APPLICANT: Manjula GANNAVARAPU  
; APPLICANT: Sebastian HOERCH  
; APPLICANT: Shubhangi KAMATKAR  
; APPLICANT: Steve G. KOVATS  
; APPLICANT: Rachel E. MEYERS  
; APPLICANT: Michael MORRISSEY  
; APPLICANT: Peter OLANDT  
; APPLICANT: Ami SEN  
; APPLICANT: Peter VEIBY  
; APPLICANT: Gordon B. MILLS  
; APPLICANT: Robert C. BAST, Jr.  
; APPLICANT: Karen LU  
; APPLICANT: Rosemarie SCHMANT  
; APPLICANT: Xumei ZHAO  
; APPLICANT: Karen GLATT  
; TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
; TITLE OF INVENTION: Assessment, Prevention, and Therapy of Ovarian Cancer  
; FILE REFERENCE: MRI-030  
; CURRENT APPLICATION NUMBER: US/10/097,340  
; CURRENT FILING DATE: 2002-03-14  
; PRIOR APPLICATION NUMBER: 60/276,025  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/325,149  
; PRIOR FILING DATE: 2001-09-26  
; PRIOR APPLICATION NUMBER: 60/276,026  
; PRIOR FILING DATE: 2001-03-14  
; PRIOR APPLICATION NUMBER: 60/324,967  
; PRIOR FILING DATE: 2001/09/26  
; PRIOR APPLICATION NUMBER: 60/311,732

;; PRIOR FILING DATE: 2001-08-10  
;; PRIOR APPLICATION NUMBER: 60/325,102  
;; PRIOR FILING DATE: 2001-09-26  
;; PRIOR APPLICATION NUMBER: 60/323,580  
;; PRIOR FILING DATE: 2001-09-19  
;; NUMBER OF SEQ ID NOS: 363  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 131  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-097-340-131

Query Match Best Local Similarity 100.0%; Score 1303; DB 14; Length 278;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLRLGLLLLLQLPAPSSASEIPKQKQAKLROREVVLDYNGMCLQSPA 60  
DB 36 MRPGPASPQRRLRLGLLLLLQLPAPSSASEIPKQKQAKLROREVVLDYNGMCLQSPA 95  
QY 61 GVPGRDGPANGIPGTGIGIPGRDGFKEGKEGECLEESFEESWTPTYKQCSMSLNYGIDL 120  
DB 96 GVPGRDGPANGIPGTGIGIPGRDGFKEGKEGECLEESFEESWTPTYKQCSMSLNYGIDL 155  
QY 121 GKIACTFTKRSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 180  
DB 156 GKIACTFTKRSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 215  
QY 181 GSPENNSTINIHRTSVSEGLCEGIGAGLVDAIIVGTCSDDYPKGDASTGMSVSR11IEE 240  
DB 216 GSPENNSTINIHRTSVSEGLCEGIGAGLVDAIIVGTCSDDYPKGDASTGMSVSR11IEE 275  
QY 241 LPK 243  
DB 276 LPK 278

## RESULT 12

US-10-177-293-496  
;; Sequence 496, Application US/10177293  
;; Publication No. US20030124128A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Lillie, James  
;; APPLICANT: Glatc, Karen  
;; APPLICANT: Zhao, Xumei  
;; APPLICANT: Gannavarpu, Manjula  
;; APPLICANT: Kamatkar, Shubhangi  
;; APPLICANT: Mertens, Maureen  
;; APPLICANT: Myer, Vic  
;; APPLICANT: Wang, Youzhen  
;; APPLICANT: Xu, Xongyao  
;; APPLICANT: Hoersch, Sebastien  
;; APPLICANT: Monahan, John  
;; APPLICANT: Meyers, Rachel E.  
;; APPLICANT: Bast Jr., Robert C.  
;; APPLICANT: Hortobagyi, Gabriel N.  
;; APPLICANT: Pusztai, Lajos  
;; APPLICANT: Meric, Funda  
;; APPLICANT: Sahin, Aysegul  
;; APPLICANT: Mills, Gordon B.  
;; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,  
;; TITLE OF INVENTION: PREVENTION, AND THERAPY OF BREAST CANCER  
;; FILE REFERENCE: MRI-038  
;; CURRENT APPLICATION NUMBER: US/10/177,293  
;; PRIOR FILING DATE: 2002-06-21  
;; PRIOR APPLICATION NUMBER: US 60/299,887  
;; PRIOR FILING DATE: 2001-06-21  
;; PRIOR APPLICATION NUMBER: US 60/301,572  
;; PRIOR FILING DATE: 2001-06-27  
;; PRIOR APPLICATION NUMBER: US 60/306,501  
;; PRIOR FILING DATE: 2001-07-18  
;; PRIOR APPLICATION NUMBER: US 60/325,002

;; PRIOR FILING DATE: 2001-09-25  
;; PRIOR APPLICATION NUMBER: US 60/362,585  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: US 60/xxx,xxx  
;; PRIOR FILING DATE: 2002-05-14  
;; NUMBER OF SEQ ID NOS: 506  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 496  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-177-293-496

Query Match Best Local Similarity 100.0%; Score 1303; DB 14; Length 278;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPGPASPQRRLRLGLLLLLQLPAPSSASEIPKQKQAKLROREVVLDYNGMCLQSPA 60  
DB 36 MRPGPASPQRRLRLGLLLLLQLPAPSSASEIPKQKQAKLROREVVLDYNGMCLQSPA 95  
QY 61 GVPGRDGPANGIPGTGIGIPGRDGFKEGKEGECLEESFEESWTPTYKQCSMSLNYGIDL 120  
DB 96 GVPGRDGPANGIPGTGIGIPGRDGFKEGKEGECLEESFEESWTPTYKQCSMSLNYGIDL 155  
QY 121 GKIACTFTKRSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 180  
DB 156 GKIACTFTKRSNSALRVLFSGSLRLKCRNACCCQWFTFNGACSGPLPEAIITYLDQ 215  
QY 181 GSPENNSTINIHRTSVSEGLCEGIGAGLVDAIIVGTCSDDYPKGDASTGMSVSR11IEE 240  
DB 216 GSPENNSTINIHRTSVSEGLCEGIGAGLVDAIIVGTCSDDYPKGDASTGMSVSR11IEE 275  
QY 241 LPK 243  
DB 276 LPK 278

## RESULT 13

US-10-301-822-77  
;; Sequence 77, Application US/10301822  
;; Publication No. US20030148410A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Millennium Pharmaceuticals, Inc.  
;; APPLICANT: Berger, Allison  
;; APPLICANT: Guillemette, Tracy L.  
;; APPLICANT: Kamatkar, Shubhangi  
;; APPLICANT: Schlegel, Robert  
;; APPLICANT: Monahan, John E.  
;; APPLICANT: Thibodeau, Stephen N.  
;; APPLICANT: Burgart, Lawrence J.  
;; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
;; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
;; TITLE OF INVENTION: THERAPY OF COLON CANCER  
;; FILE REFERENCE: MP001-029P2RNM  
;; CURRENT APPLICATION NUMBER: US/10/301,822  
;; PRIOR FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US 60/339,971  
;; PRIOR FILING DATE: 2001-12-10  
;; PRIOR APPLICATION NUMBER: US 60/361,978  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: US 60/381,988  
;; PRIOR FILING DATE: 2002-05-20  
;; NUMBER OF SEQ ID NOS: 228  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 77  
;; LENGTH: 278  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-301-822-77

Query Match Best Local Similarity 100.0%; Score 1303; DB 14; Length 278;  
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRPOGPASFORLRLGLLLLLLLQLPAPSSASEIPKQKQAKLRQREVVLDLYNGMCLQGPA 60
DB 36 MRPOGPASFORLRLGLLLLLLLQLPAPSSASEIPKQKQAKLRQREVVLDLYNGMCLQGPA 95
QY 61 GVGGRDGSFGANGIPGTGPIGRDGFKEGKEGCELRSEFESWTPNTYKQCSWSLNYGIDL 120
DB 96 GVGGRDGSFGANGIPGTGPIGRDGFKEGKEGCELRSEFESWTPNTYKQCSWSLNYGIDL 155
QY 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLDQ 180
DB 156 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLDQ 215
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
DB 216 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 275
QY 241 LPK 243
DB 276 LPK 278
```

## RESULT 14

```
US-10-296-115-1261
; Sequence 1261, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hygeed Inc.
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25
; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 1261
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-296-115-1261
```

Query Match 100.0%; Score 1303; DB 15; Length 278;

Best Local Similarity 100.0%; Pred. No. 1.8e-115;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRPOGPASFORLRLGLLLLLLLQLPAPSSASEIPKQKQAKLRQREVVLDLYNGMCLQGPA 60
DB 36 MRPOGPASFORLRLGLLLLLLLQLPAPSSASEIPKQKQAKLRQREVVLDLYNGMCLQGPA 95
QY 61 GVGGRDGSFGANGIPGTGPIGRDGFKEGKEGCELRSEFESWTPNTYKQCSWSLNYGIDL 120
DB 96 GVGGRDGSFGANGIPGTGPIGRDGFKEGKEGCELRSEFESWTPNTYKQCSWSLNYGIDL 155
QY 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLDQ 180
DB 156 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLDQ 215
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
DB 216 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 275
QY 241 LPK 243
DB 276 LPK 278
```

GENERAL INFORMATION:

```
; APPLICANT: Endege, Wilson O.
; APPLICANT: Ford, Donna
; APPLICANT: Ganavathapu, Manjula
; APPLICANT: Glat, Karen
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamakar, Shubhangi
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Xu, Yong Yao
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND PROTEINS FOR
; TITLE OF INVENTION: THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; FILE REFERENCE: MRI-065
; CURRENT APPLICATION NUMBER: US/10/961,139
; PRIOR FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: 60/509,171
; PRIOR FILING DATE: 2003-10-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-961-139-2
```

Query Match 100.0%; Score 1303; DB 18; Length 278;

Best Local Similarity 100.0%; Pred. No. 1.8e-115;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MRPOGPASFORLRLGLLLLLLLQLPAPSSASEIPKQKQAKLRQREVVLDLYNGMCLQGPA 60
DB 36 MRPOGPASFORLRLGLLLLLLLQLPAPSSASEIPKQKQAKLRQREVVLDLYNGMCLQGPA 95
QY 61 GVGGRDGSFGANGIPGTGPIGRDGFKEGKEGCELRSEFESWTPNTYKQCSWSLNYGIDL 120
DB 96 GVGGRDGSFGANGIPGTGPIGRDGFKEGKEGCELRSEFESWTPNTYKQCSWSLNYGIDL 155
QY 121 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLDQ 180
DB 156 GKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGGLPIEAIYYLDQ 215
QY 181 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 240
DB 216 GSPENNSTINIHRTSSVEGLCEGIGAGLVDAIWMGTCSDPYKGDASTGMNSVSRITIEE 275
QY 241 LPK 243
DB 276 LPK 278
```

Search completed: August 4, 2005, 04:45:27

Job time : 87.0824 secs

RESULT 15  
US-10-961-139-2  
; Sequence 2, Application US/10961139  
; Publication No. US2005015313A1

This Page Blank (uspio)

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:56 ; Search time 24.6601 Seconds  
(without alignments) 741.643 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307  
Sequence: 1 MHQGRASPOLLLGLFLV.....GDASTGMSVSRILIEBLK 245

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/1aa/5A COMB.pep.\*  
2: /cgn2\_6/ptodata/1/1aa/5A COMB.pep.\*  
3: /cgn2\_6/ptodata/1/1aa/6A COMB.pep.\*  
4: /cgn2\_6/ptodata/1/1aa/6A COMB.pep.\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS COMB.pep.\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1307	100.0	245	4	US-09-692-081-2
2	1307	100.0	277	4	US-09-692-081-5
3	1225	93.7	243	4	US-09-692-081-4
4	1222	93.5	243	4	US-09-489-847-205
5	1216	93.0	243	4	US-09-834-759-514
6	1216	93.0	278	4	US-09-834-759-514
7	1101	84.2	276	4	US-09-205-258-958
8	968	74.1	197	4	US-09-834-759-516
9	968	74.1	232	4	US-09-834-759-517
10	353	27.0	66	4	US-09-205-258-962
11	296	22.6	52	4	US-09-205-258-961
12	270	20.7	51	4	US-09-205-258-963
13	248	19.0	46	4	US-09-834-759-518
14	215	16.4	52	4	US-09-205-258-960
15	141	10.8	26	4	US-09-834-759-519
16	130	9.9	93	4	US-09-489-847-240
17	130	9.9	93	4	US-09-489-847-362
18	124.5	9.5	1516	4	US-09-949-016-8209
19	121	9.3	128	3	US-09-227-357-190
20	121	9.3	623	3	US-09-029-348-3
21	120	9.2	755	4	US-09-029-348-2
22	119.5	9.1	246	2	US-08-463-911-4
23	119.5	9.1	313	4	US-09-949-016-9265
25	119.5	9.1	1670	4	US-09-949-016-5883
26	117	9.0	1057	3	US-08-931-820-4
27	117	9.0	1078	3	US-08-963-825-21

28	117	9.0	1078	3	US-09-500-811-21	Sequence 21, Appl
29	117	9.0	1078	3	US-09-570-573-21	Sequence 21, Appl
30	117	9.0	1078	3	US-09-548-608-21	Sequence 21, Appl
31	116	8.9	258	4	US-09-976-594-815	Sequence 815, App
32	116	8.9	744	4	US-09-949-016-9607	Sequence 9607, Ap
33	115.5	8.8	349	6	5510466-2	Patent No. 5510466
34	115.5	8.8	349	6	5510466-2	Patent No. 5510466
35	115.5	8.8	453	6	5510466-4	Patent No. 5510466
36	115.5	8.8	453	6	5510466-4	Patent No. 5510466
37	115.5	8.8	684	3	US-08-555-669-12	Sequence 12, Appl
38	115.5	8.8	684	3	US-09-073-663-12	Sequence 12, Appl
39	115	8.8	245	4	US-09-552-225A-4	Sequence 4, Appl
40	115	8.8	245	4	US-09-552-204A-4	Sequence 4, Appl
41	115	8.8	1078	4	US-09-949-016-1185	Sequence 1185, A
42	115	8.8	1143	4	US-09-949-016-6137	Sequence 6137, Ap
43	114.5	8.8	248	4	US-09-600-932-28	Sequence 28, Appl
44	114	8.7	310	3	US-09-219-849-47	Sequence 47, Appl
45	114	8.7	451	1	US-08-453-117-2	Sequence 2, Appl

#### ALIGNMENTS

```
RESULT 1
US-09-692-081-2
; Sequence 2, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-692-081-2

Query Match      100.0%; Score 1307; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 2.2e-133; Indels 0; Gaps 0;
Matches 245; Conservative 0; Mismatches 0;

QY      1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPKVKOKALIROREVDLYNGMCLQG 60
Db      1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPKVKOKALIROREVDLYNGMCLQG 60

QY      61 PAGVPRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPNPKQCSWSSLYNGI 120
Db      61 PAGVPRDGSFGANGIPGTGIPGRDGFKEGKECLRESFEESWTPNPKQCSWSSLYNGI 120

QY      121 DKGIAECFTMRNSALRVFSGSLRLKCNACCORRYTFNAGCGPLPIEATLYL 180
Db      121 DKGIAECFTMRNSALRVFSGSLRLKCNACCORRYTFNAGCGPLPIEATLYL 180

QY      181 DQGSBELNSTINIHRTSSVEGLCEGIGAGLVVAIIVGTCSDPYPKDASTGMSVSRILI 240
Db      181 DQGSBELNSTINIHRTSSVEGLCEGIGAGLVVAIIVGTCSDPYPKDASTGMSVSRILI 240

QY      241 EELPK 245
Db      241 EELPK 245

RESULT 2
US-09-692-081-5
; Sequence 5, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
```

```
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; CURRENT FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-692-081-5
```

```
Query Match          100.0%; Score 1307; DB 4; Length 277;
Best Local Similarity 100.0%; Pred. No. 2.6e-133;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKALIRQREVVDLYNMGCLQG 60
DB 33 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKALIRQREVVDLYNMGCLQG 92
QY 61 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTNTNYKQCSWSLNYGI 120
DB 93 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTNTNYKQCSWSLNYGI 152
QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCOMWFTFNGACSGPLPIEAIIVL 180
DB 153 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCOMWFTFNGACSGPLPIEAIIVL 212
QY 181 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGNSVSR111 240
DB 213 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGNSVSR111 272
QY 241 EELPK 245
DB 273 EELPK 277
```

## RESULT 3

```
US-09-692-081-4
; Sequence 4, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; CURRENT FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-692-081-4
```

```
Query Match          93.7%; Score 1225; DB 4; Length 243;
Best Local Similarity 94.7%; Pred. No. 1.7e-124;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKALIRQREVVDLYNMGCLQG 60
DB 1 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKALIRQREVVDLYNMGCLQG 58
QY 61 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTNTNYKQCSWSLNYGI 120
DB 59 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTNTNYKQCSWSLNYGI 118
QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCOMWFTFNGACSGPLPIEAIIVL 180
DB 119 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCOMWFTFNGACSGPLPIEAIIVL 178
```

```
QY 181 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGNSVSR111 240
DB 179 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGNSVSR111 238
QY 241 EELPK 245
DB 239 EELPK 243
```

## RESULT 4

```
US-09-489-847-205
; Sequence 205, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 205
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-205
```

```
Query Match          93.5%; Score 1222; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 3.6e-124;
Matches 231; Conservative 4; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKALIRQREVVDLYNMGCLQG 60
DB 1 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKALIRQREVVDLYNMGCLQG 58
QY 61 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTNTNYKQCSWSLNYGI 120
DB 59 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECLEESFEESWTNTNYKQCSWSLNYGI 118
QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCOMWFTFNGACSGPLPIEAIIVL 180
DB 119 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCOMWFTFNGACSGPLPIEAIIVL 178
QY 181 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGNSVSR111 240
DB 179 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPYPKGDASTGNSVSR111 238
QY 241 EELPK 245
DB 239 EELPK 243
```

## RESULT 5

```
US-09-834-759-514
; Sequence 514, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitchell, Jennifer L.
; APPLICANT: Xu, Jiaqichun
```

```

; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ. ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 514
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-514
```

```

Query Match      93.0%; Score 1216; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 1.6e-123;
Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
```

```

QY 1 MHQGRASPOLLLGLFVLLLLQLSPSSASENPVKKALIRQREVVDLYNGCTCLOG 60
    |||
    1 MRQGRASPOLLLGLFVLLLLQLSPSSASENPVKKALIRQREVVDLYNGCTCLOG 58
DB 59 PAVPGRDSSPGANVIGPTGIPGRDGFKEGKECELRSEFESWTPNYKOCSSSLNYGI 118
    |||
QY 121 DLGKIAECTFTKRRNSALRVLPFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYVL 180
    |||
DB 119 DLGKIAECTFTKRRNSALRVLPFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYVL 178
QY 181 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNNVSRIII 240
    |||
DB 179 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNNVSRIII 238
QY 241 EELPK 245
    |||
DB 239 EELPK 243
```

```

RESULT 6
US-09-834-759-515
; Sequence 515, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ. ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515
```

```

Query Match      93.0%; Score 1216; DB 4; Length 278;
Best Local Similarity 94.3%; Pred. No. 2e-123;
Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
```

```

QY 1 MHQGRASPOLLLGLFVLLLLQLSPSSASENPVKKALIRQREVVDLYNGCTCLOG 60
    |||
DB 36 MRQGRASPOLLLGLFVLLLLQLSPSSASENPVKKALIRQREVVDLYNGCTCLOG 93
```

```

QY 61 PAVPGRDSSPGANVIGPTGIPGRDGFKEGKECELRSEFESWTPNYKOCSSSLNYGI 120
    |||
DB 94 PAVPGRDSSPGANVIGPTGIPGRDGFKEGKECELRSEFESWTPNYKOCSSSLNYGI 153
QY 121 DLGKIAECTFTKRRNSALRVLPFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYVL 180
    |||
DB 154 DLGKIAECTFTKRRNSALRVLPFSGSLRLKCRNACCORWYFTFNGACSGPLPIEAIYVL 213
QY 181 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNNVSRIII 240
    |||
DB 214 DQSSPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGNNVSRIII 273
QY 241 EELPK 245
    |||
DB 274 EELPK 278
```

```

RESULT 7
US-09-205-258-958
; Sequence 958, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 958
; LENGTH: 276
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-958
```

```
Query Match      84.2%; Score 1101; DB 4; Length 276;
Best Local Similarity 98.5%; Pred. No. 5.7e-111;
Matches 202; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
```

```
QY 41 KALIRREVDLYNGMCLCGPAGVPRGDSRGANGIPGTGIGRGGFKGKCECLRESEF 100
DB 72 KRIRIREVVDLYNGMCLCGPAGVPRGDSRGANGIPGTGIGRGGFKGKCECLRESEF 131
QY 101 EESWTPNYKCCSSSLNYGIDLCIAECTFKRSNSALRVLPFSGSLRLKCRNACCOMRY 160
DB 132 EESWTPNYKCCSSSLNYGIDLCIAECTFKRSNSALRVLPFSGSLRLKCRNACCOMRY 191
QY 161 FTENGACSGPLIEAIIYVDGSPSELNSTINIHRTSSVGLCEGIGAGIIVDAIVWGTC 220
DB 192 FTENGACSGPLIEAIIYVDGSPSELNSTINIHRTSSVGLCEGIGAGIIVDAIVWGTC 251
QY 221 SDYKGDASTGMSVSRIITIEELPK 245
DB 252 SDYKGDASTGMSVSRIITIEELPK 276
```

```
RESULT 8
US-09-834-759-516
; Sequence 516, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
```

```
; APPLICANT: JIANG, YUGU
; APPLICANT: DILLON, DAVID C.
; APPLICANT: MITCHAM, JENNIFER L.
; APPLICANT: XU, JIANGCHUN
; APPLICANT: HARLOCKER, SUSAN L.
; APPLICANT: HEPLER, WILLIAM T.
; APPLICANT: HENDERSON, ROBERT A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 516
; LENGTH: 197
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-834-759-516
```

```
Query Match      74.1%; Score 968; DB 4; Length 197;
Best Local Similarity 93.0%; Pred. No. 9.3e-97;
Matches 185; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
```

```
QY 1 MHPGRRASPQILLGLFVLLILLQLSAPSSASENPKYKQKALIRQREVVDLYNGMCLCG 60
DB 1 MHPGRRASPQIRG--LILLILLQLPAPSSASEIPKQKQKQRLQREVVDLYNGMCLCG 58
QY 61 PAVGPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPNYKCCSSSLNYGI 120
DB 59 PAVGPGRDGPANGIPGTGIPGRDGFKEGKECLRESFEESWTPNYKCCSSSLNYGI 118
QY 121 DLGIAECTFKRSNSALRVLPFSGSLRLKCRNACCOMRYFPNAGCSGPLIEAIIYL 180
DB 119 DLGIAECTFKRSNSALRVLPFSGSLRLKCRNACCOMRYFPNAGCSGPLIEAIIYL 178
QY 181 DQGSPELNSTINIHRTSSV 199
DB 179 DQGSPELNSTINIHRTSSV 197
```

```
RESULT 9
US-09-834-759-517
; Sequence 517, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: JIANG, YUGU
; APPLICANT: DILLON, DAVID C.
; APPLICANT: MITCHAM, JENNIFER L.
; APPLICANT: XU, JIANGCHUN
; APPLICANT: HARLOCKER, SUSAN L.
; APPLICANT: HEPLER, WILLIAM T.
; APPLICANT: HENDERSON, ROBERT A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 517
; LENGTH: 232
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-834-759-517
```

```
Query Match      74.1%; Score 968; DB 4; Length 232;
Best Local Similarity 93.0%; Pred. No. 1.2e-96;
Matches 185; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
```

```
QY 1 MHPGRRASPQILLGLFVLLILLQLSAPSSASENPKYKQKALIRQREVVDLYNGMCLCG 60
DB 36 MHPGRRASPQIRG--LILLILLQLPAPSSASEIPKQKQKQRLQREVVDLYNGMCLCG 93
```



Qy	6	PAGPGRDSDSGANGLPGRIPIGRPRGFGEKGECLARESPESMTPTNPKOCSMSLNYGI	120
Db	94	PAGPGRDSDSGANVLPGTPIPIGRGFGEKGECLARESPESMTPTNPKOCSMSLNYGI	153
Qy	121	DLGKIACEFTTKRRSNAIRVLVFGSGLRLKCNACCOMWYFTFNGAECSGPIPEALITYL	180
Db	154	DLGKIACEFTTKRRSNAIRVLVFGSGLRLKCNACCOMWYFTFNGAECSGPIPEALITYL	213
Qy	181	DOGSPELNSTINIHRTSSV	199
Db	214	DOGSPELNSTINIHRTSSV	232
RESULT 10			
US-09-205-258-962			
Sequence 962, Application US/09205258			
Patent No. 6525174			
GENERAL INFORMATION:			
APPLICANT: Young et al.			
TITLE OF INVENTION: 207 Human Secreted Proteins			
FILE REFERENCE: P2007P1			
CURRENT APPLICATION NUMBER: US/09/205,258			
CURRENT FILING DATE: 1998-12-04			
EARLIER APPLICATION NUMBER: PCT/US98/11422			
EARLIER FILING DATE: 1998-06-04			
EARLIER APPLICATION NUMBER: 60/048,885			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,375			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,881			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,880			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,896			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,020			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,876			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,895			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,884			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,894			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,971			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,964			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,882			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,899			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,893			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,900			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,901			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,892			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,915			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,019			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,970			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,972			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/048,916			
EARLIER FILING DATE: 1997-06-06			
EARLIER APPLICATION NUMBER: 60/049,373			

```

1  EARLIER FILING DATE: 1997-06-06
2  EARLIER APPLICATION NUMBER: 60/048,875
3  EARLIER FILING DATE: 1997-06-06
4  EARLIER APPLICATION NUMBER: 60/049,374
5  EARLIER FILING DATE: 1997-06-06
6  EARLIER APPLICATION NUMBER: 60/048,917
7  EARLIER FILING DATE: 1997-06-06
8  EARLIER APPLICATION NUMBER: 60/048,949
9  EARLIER FILING DATE: 1997-06-06
10 EARLIER APPLICATION NUMBER: 60/048,974
11 EARLIER FILING DATE: 1997-06-06
12 EARLIER APPLICATION NUMBER: 60/048,983
13 EARLIER FILING DATE: 1997-06-06
14 EARLIER APPLICATION NUMBER: 60/048,897
15 EARLIER FILING DATE: 1997-06-06
16 EARLIER APPLICATION NUMBER: 60/048,898
17 EARLIER FILING DATE: 1997-06-06
18 EARLIER APPLICATION NUMBER: 60/048,962
19 EARLIER FILING DATE: 1997-06-06
20 EARLIER APPLICATION NUMBER: 60/048,963
21 EARLIER FILING DATE: 1997-06-06
22 EARLIER APPLICATION NUMBER: 60/048,877
23 EARLIER FILING DATE: 1997-06-06
24 EARLIER APPLICATION NUMBER: 60/048,878
25 EARLIER FILING DATE: 1997-06-06
26 EARLIER APPLICATION NUMBER: 60/070,923
27 EARLIER FILING DATE: 1997-12-18
28 EARLIER APPLICATION NUMBER: 60/092,921
29 EARLIER FILING DATE: 1998-07-15
30 EARLIER APPLICATION NUMBER: 60/094,657
31 EARLIER FILING DATE: 1998-07-30
32 NUMBER OF SEQ ID NOS: 1227
33 SOFTWARE: PatentIn Ver. 2.0
34 SEQ ID NO 962
35 LENGTH: 66
36 TYPE: PR1
37 ORGANISM: Homo sapiens
38 US-09-205-258-962
39
40 Query Match          27.0%; Score 353; DB 4; Length 66;
41 Best Local Similarity 98.5%; Pred. No. 9,3e-31;
42 Matches 65; Conservative 1; Mismatches 0; Indels 0; Gaps 0
43
44 Qy      130  FTTGRSNSALRVLFSGSLRLKCRNACCRWYFTFNGAECGSPDIEAIIYIDGSSPLNS 1699
45      1  FTTGRSNSALRVLFSGSLRLKCRNACCRWYFTFNGAECGSPDIEAIIYIDGSSPLNS 60
46
47 Db
48      190  TINIR 195
49      |||||
50      61  TINIR 66
51
52 RESULT 11
53 US-09-205-258-961
54 Sequence 961, Application US/09205258
55 Patent No. 6525174
56 GENERAL INFORMATION:
57 APPLICANT: Young et al.
58 TITLE OF INVENTION: 207 Human Secreted Proteins
59 FILE REFERENCE: P2007P1
60 CURRENT APPLICATION NUMBER: US/09/205,258
61 CURRENT FILING DATE: 1998-12-04
62 EARLIER APPLICATION NUMBER: PCT/US98/11422
63 EARLIER FILING DATE: 1998-06-04
64 EARLIER APPLICATION NUMBER: 60/048,885
65 EARLIER FILING DATE: 1997-06-06
66 EARLIER APPLICATION NUMBER: 60/049,375
67 EARLIER FILING DATE: 1997-06-06
68 EARLIER APPLICATION NUMBER: 60/048,881
69 EARLIER FILING DATE: 1997-06-06
70 EARLIER APPLICATION NUMBER: 60/048,880
71 EARLIER FILING DATE: 1997-06-06
72 EARLIER APPLICATION NUMBER: 60/048,896

```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 961
; LENGTH: 52
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-205-258-961
;
; Query Match 22.6%; Score 296; DB 4; Length 52;
; Best Local Similarity 100.0%; Pred. No. 1e-24;
; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 79 TGTGTRDGFKEKGCRCRESFEESWTPTNYKCCSSINLYGIDLGKIACTP 130
; DB 1 TGTGTRDGFKEKGCRCRESFEESWTPTNYKCCSSINLYGIDLGKIACTP 52
;
; RESULT 12
; US-09-205-258-963
; Sequence 963, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
```

```

; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 963
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-963

```

```

Query Match      20.7%; Score 270; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 6,4e-22;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      195 RTSSVGLCGIGAGLVDAIWMGTCSDPKGDASTGMNSVSRITIELPK 245
Db      1 RTSSVGLCGIGAGLVDAIWMGTCSDPKGDASTGMNSVSRITIELPK 51

```

```

RESULT 13
US-09-834-759-518
; Sequence 518, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121,470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 518
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-518
Query Match      19.0%; Score 248; DB 4; Length 46;

```

```

Best Local Similarity 100.0%; Pred. No. 1.3e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      200 EGIICGIGAGLVDAIWMGTCSDPKGDASTGMNSVSRITIELPK 245
Db      1 EGIICGIGAGLVDAIWMGTCSDPKGDASTGMNSVSRITIELPK 46

```

```

RESULT 14
US-09-205-258-960
; Sequence 960, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917

```

EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30  
NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 960  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-205-258-960

Query Match 16.4%; Score 215; DB 4; Length 52;  
Best Local Similarity 95.2%; Pred. No. 6.1e-16;  
Matches 40; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 41 KALIRORVVDLYNGMCLOGPAGVPGRDSPGANGIPGTPI 82  
DB 11 KRIROREVVDLYNGMCLOGPAGVPGRDSPGANGIPGTPI 52

RESULT 15  
US-09-834-759-519  
Sequence 519, Application US/09834759  
Patent No. 6680197  
GENERAL INFORMATION:  
APPLICANT: Jiang, Yugu  
APPLICANT: Dillon, Devin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Xu, Jiangchun  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Hepler, William T.  
APPLICANT: Henderson, Robert A.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.470C9  
CURRENT APPLICATION NUMBER: US/09/834,759  
CURRENT FILING DATE: 2001-04-13  
NUMBER OF SEQ ID NOS: 547  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 519  
LENGTH: 26  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-834-759-519

Query Match 10.8%; Score 141; DB 4; Length 26;  
Best Local Similarity 100.0%; Pred. No. 2.4e-08;  
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 220 CSDYPKGDASTGMNSVSRRIIEELPK 245  
|||||

DB 1 CSDYPKGDASTGMNSVSRRIIEELPK 26  
Search completed: August 4, 2005, 04:23:52  
Job time : 25.6601 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:16 ; Search time 19.5359 Seconds

(without alignments)  
1206.653 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307

Sequence: 1 MHQGRAPAPQLLGFLVVL.....GDASTGMSVSRRIIEELPK 245

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR 79: \*  
2: PIR 79: \*  
3: PIR 79: \*  
4: PIR 79: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	127.5	9.8	289	T20177	hypothetical prote
2	125.5	9.6	298	T27644	hypothetical prote
3	125.5	9.6	1752	A45407	collagen alpha 3(I
4	123.5	9.4	311	T15268	hypothetical prote
5	122.5	9.4	920	A45748	collagen alpha 1(V
6	122.5	9.4	1549	T48103	type VII collagen
7	122	9.3	330	S46557	collagen alpha 1(X
8	121	9.3	178	A39762	collagen alpha 1(X
9	121	9.3	1466	CGHUTL	collagen alpha 1(I
10	121	9.3	1763	SGHUTL	collagen alpha 2(I
11	119.5	9.1	246	S29328	complement subcomp
12	119.5	9.1	310	T29731	hypothetical prote
13	119.5	9.1	323	A61396	collagen alpha 1(I
14	119.5	9.1	1670	CGH03B	collagen alpha 3(I
15	119.5	9.1	2944	A54849	collagen alpha 1(V
16	119	9.1	298	JC1448	collagen col-34 -
17	119	9.1	299	T29956	hypothetical prote
18	119	9.1	1049	CGH07S	collagen alpha 1(I
19	118	9.0	468	T31631	hypothetical prote
20	118	9.0	1669	CGMS4B	collagen alpha 1(I
21	118	9.0	1744	S41920	collagen alpha 1(I
22	117.5	9.0	324	A31920	collagen sct-1 pre
23	117.5	9.0	673	CGBO6C	collagen alpha 1(I
24	117	9.0	253	S49158	complement protein
25	116.5	8.9	58	A45526	collagen homolog A
26	116.5	8.9	300	T24482	hypothetical prote
27	116	8.9	289	T26812	hypothetical prote
28	116	8.9	302	T15936	hypothetical prote
29	116	8.9	744	S15435	collagen alpha 1(V

30	115.5	8.8	349	2	S08278	macrophage scaveng
31	115.5	8.8	428	2	T24769	hypothetical prote
32	115.5	8.8	453	2	S08276	macrophage scaveng
33	115	8.8	245	1	CHUQC	complement subcomp
34	115	8.8	296	2	A31219	collagen 1 - Caeno
35	115	8.8	301	2	B31219	collagen 2 - Caeno
36	115	8.8	744	1	A34246	collagen alpha 1(V
37	115	8.8	744	1	S23298	collagen alpha 1(V
38	115	8.8	1042	1	CGCH1S	collagen alpha 1(I
39	115	8.8	1142	2	JX0369	collagen alpha 1(X
40	114.5	8.8	248	1	LNHUP1	pulmonary surfacta
41	114.5	8.8	1496	1	CGH02V	collagen alpha 2(V
42	114.5	8.8	1497	2	I49607	procollagen type V
43	114	8.7	327	2	T29031	hypothetical prote
44	114	8.7	451	2	A38415	macrophage scaveng
45	114	8.7	671	1	CGRT1S	collagen alpha 1(I

## ALIGNMENTS

## RESULT 1

T20177  
hypothetical protein C53B4.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #ext\_change 09-Jul-2004

C:Accession: T20177

R:Berke, M.  
submitted to the EMBL Data Library, December 1995

A:Reference number: Z19233

A:Accession: T20177

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-289 <MIL>

A:Cross-references: UNIPROT:Q18799; EMBL:Z68215; PIDN:CAA92453.1; GSPDB:GN00022; CESP:C.

A:Experimental source: clone C53B4

C:Genetics:

A:Gene: CESP:C53B4.5

A:Map position: 4

Query Match

Best Local Similarity 9.8%; Score 127.5; DB 2; Length 289;

Matches 25; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

QY

59 QGPAGVPRGDS-----PGANGTPTGTPGRDGFGEKGC 95

DB 220 RGPAGQPGHDAQCGPGEKANGBPQGRDGPGRPGQGRDGHGEKVC 271

## RESULT 2

T27644  
hypothetical protein ZK1010.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #ext\_change 09-Jul-2004

C:Accession: T27644

R:Gardner, A.  
submitted to the EMBL Data Library, November 1996

A:Reference number: Z20398

A:Accession: T27644

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-298 <MIL>

A:Cross-references: UNIPROT:Q18286; EMBL:Z62083; PIDN:CAB04973.1; GSPDB:GN00021; CESP:ZK

A:Experimental source: clone ZK1010

C:Genetics:

A:Gene: CESP:ZK1010.7

A:Map position: 3

A:Introns: 54/3; 236/1

Query Match

Best Local Similarity 9.6%; Score 125.5; DB 2; Length 298;

Matches 30; Conservative 6; Mismatches 29; Indels 11; Gaps 2;



```

QY      60 GPAGVPGR---DSPGANGIPGTPGTPGRDGFKEGKE 94
          |||||: ||||| ||||| ||||| |||||
Db      426 GPPGVPGKTGEDGKPLNGKNGEPGDPGEDGRKEGKD 466

```

## RESULT 7

collagen alpha 1(XIV) chain - human (fragments)  
C/Species: Homo sapiens (man)  
C/Date: 15-Jul-1995 #sequence revision 03-Nov-1995 #text\_change 09-Jul-2004  
C/Accession: S46657, S37550, F44398  
R/Brown, J.C.; Golblik, R.; Mann, K.; Timpl, R.  
Matrix Biol. 14, 287-295, 1994  
A/Title: Structure and stability of the triple-helical domains of human collagen XIV.  
A/Reference number: S46657; MUID:95128518; PMID:783751  
A/Accession: S46657  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 26-330 <BRO>  
A/Cross-references: UNIPROT:Q9UDF6  
R/Brown, J.C.; Golblik, R.; Mann, K.; Timpl, R.  
submitted to the EMBL Data Library, October 1993  
A/Description: Structure and stability of the triple helical domains of human collagen XIV.  
A/Reference number: S37749  
A/Accession: S37750  
A/Molecule type: protein  
A/Residues: 203-330 <BR2>  
R/Brown, J.C.; Mann, K.; Wiedemann, H.; Timpl, R.

```

Query Match      9.3%; Score 122; DB 2; Length 330;
Best Local Similarity 31.9%; Pred. No. 0.0019;
Matches 36; Conservative 10; Mismatches 33; Indels 34; Gaps 5

QY      3  PGRRAAPQLLLGLFLVLLLLQLQAPSSASNPVKYKALIR--QRQYVD-----L 52
Db      155 PQAALCPPEV-----PQAAGGRRGGDLQQAANVRYARVAKYEQILQISMAR 201

```

Dy 53 YNGM-----CLGGPAGVPGRDGSPGA---NGIGTGPGLPSGDGFKEGEKGS 94  
|| : ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |  
Db 202 YNQIPRSHSSIRTVQGRPFGEGRPGSRPAFGEGRPCTPGFPGAAGVPRTPEE 25

## RESULT 8

collagen alpha 1(XIV) chain - bovine (fragments)  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 14-Feb-1992 #sequence\_revision 05-Apr-1995 #ext\_change 09-Jul-2004  
C/Accession: A39762; S23815; S16905  
R/Dublet, B.; van der Rest, M.  
J. Biol. Chem. 266, 6853-6858, 1991  
A/Title: Type XIV collagen; a new homotrimeric molecule extracted from fetal bovine skin  
R/Reference number: A39762; MUID:51201330; PMID:21261301

```

A:Accession: A39762
A:Molecule type: protein
A:Residues: 1-19;20-61;62-71;72-81;82-123;124-178 <DUB>
A:Cross-references: UNIPROT:Q7M2M9
R:Auberer-Fouchier, E.; Font, B.; Eichenberger, D.; Goldschmidt, D.; Lethias, C.; van der
J. Biol. Chem. 267, 15759-15764, 1992
A:Title: Purification and characterization of native type XIV collagen.
A:Reference number: S23815; MUID:92348439; PMID:1322405
A:Accession: S23815
A:Molecule type: protein
A:Residues: 'XR',82-94,'X',96 <AUB>
A:Complex: homotrimer
C:Keywords: coiled coil; disulfide bond; extracellular matrix; homotrimer; hydroxylysine
Query Match 9.3%; Score 121; DB 2; Length 178;
Best Local Similarity 33.0%; Pred. No. 0.0012;
Matches 32; Conservative 7; Mismatches 34; Indels 24; Gaps . 4;

```

QY 61 PAVPGRDGS PGA---NGIPGTPIGRDGFKEGKE 94  
 86 PPGEPGRPGSPGT PGEQGPPGAGFPGNAGVPGSXGE 122  
 DB

## RESULT 5

collagen alpha1(III) chain precursor - human  
N/Alternate names: procollagen alpha 1 (III) chain  
C/Species: Homo sapiens (man)  
C/Date: 24-Apr-1994 #sequence\_revision 01-Sep-1995 #text\_change 09-Jul-2004  
C/Accession: S05272, S04642, PE0011, S01726, S04887, A50399, A94562, I51868, S59511, A9  
R/Prockop, D.J. submitted to the EMBL Data Library, February 1989  
A/Reference number: S05272  
A/Accession: S05272  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-1240, 'V', 1242-1466 <PRC>  
A/Cross-references: UNIPROT:P02461; EMBL:X14420, NID:g30057, PIDN:CAA32583.1, PID:g3005  
R/Ala-Kokko, L.; Kontusarari, S.; Baldwin, C.T.; Kuivanen, H.; Prockop, D.J.  
Biochem. J. 260, 509-516, 1989  
A/Title: Structure of cDNA clones coding for the entire prepro-alpha1(III) chain of hum  
ences.  
A/Reference number: S04642, MUID:89350838, PMID:2764886  
A/Accession: S04642  
A/Molecule type: mRNA  
A/Residues: 1-1196 <ALA>  
A/Cross-references: EMBL:X14420, NID:g30057, PIDN:CAA32583.1, PID:g30058  
A/Note: the complete sequence is not shown  
R/Benson-Chanda, V.; Su, M.W.; Weil, D.; Chu, M.L.; Ramirez, F.  
Gene 78, 255-265, 1989  
A/Title: Cloning and analysis of the 5' portion of the human type-III procollagen gene  
A/Reference number: PE0011, MUID:89378752, PMID:2777083  
A/Accession: PE0011  
A/Molecule type: DNA  
A/Residues: 1-176 <BEN>  
A/Cross-references: GB:Z62939, NID:g180813, PIDN:AAA52040.1, PID:g180814  
R/Toman, P.D.; Ricca, G.A.; de Crombrughe, B.  
Nucleic Acids Res. 16, 7201, 1988  
A/Title: Nucleotide sequence of a cDNA coding for the amino-terminal region of human pr  
A/Reference number: S01726, MUID:88303360, PMID:3405773  
A/Accession: S01726  
A/Molecule type: mRNA  
A/Residues: 1-170 <TM>  
A/Cross-references: EMBL:X07240, NID:g30060, PIDN:CAA30229.1, PID:g30061  
R/Janecko, R.A.; Ramirez, F.  
Nucleic Acids Res. 17, 6742, 1989  
A/Title: Nucleotide and amino acid sequences of the entire human alpha-1 (III) collagen  
A/Reference number: S04887, MUID:89386015, PMID:2780304

A/Accession: S04887  
 A/Molecule type: mRNA  
 A/Residues: 149-163,'G',164-240,'D',242-471,'D',473-487,'L',489,'S',491-613,'Y',615-634,  
 A/Cross-references: EMBL:X15332; NID:g29545; PIDD:CAA33387.1; PID:g930045  
 A/Note: the authors' translation of residues 905-932 is inconsistent with the nucleotide  
 R/Seyer, J.M.; Kang, A.H.  
 Biochemistry 16, 1158-1164, 1977  
 A/Title: Covalent structure of collagen: amino acid sequence of cyanogen bromide peptide  
 A/Reference number: A90399; MUID:77134724; PMID:557335  
 A/Accession: A90399  
 A/Molecule type: protein  
 A/Residues: 'V',169-225,229-232,'P',234-292,'D',294-398 <SEY1>  
 A/Experimental source: liver  
 A/Note: sequence corrected by A94562; attachment of 2-O-alpha-D-glucosyl-O-beta-D-galact  
 R/Seyer, J.M.  
 submitted to the Atlas, December 1977  
 A/Reference number: A94562  
 A/Accession: A94562  
 A/Molecule type: protein  
 A/Residues: 'V',169-225,229-277,'A',279-292,'D',294,'S',296-398 <SEY2>  
 A/Experimental source: liver  
 A/Note: author submitted corrections to A90399  
 R/Milewicz, D.M.; Witz, A.M.; Smith, A.C.; Manchester, D.K.; Waldstein, G.; Byers, P.H.  
 Am. J. Hum. Genet. 53, 62-70, 1993  
 A/Title: Parental somatic and germ-line mosaicism for a multiexon deletion with unusual  
 splicing.  
 A/Reference number: 151868; MUID:93304430; PMID:8317500  
 A/Accession: 151868  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 186-194 <ML>  
 A/Cross-references: GB:S62925; NID:g386425; PIDD:AAJ3937.1; PID:g4261637  
 R/Chiodo, A.A.; Silience, D.O.; Cole, W.G.; Bateman, J.F.  
 Biochem. J. 311, 939-943, 1995  
 A/Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3  
 A/Reference number: S59511; MUID:96067614; PMID:7487954  
 A/Accession: S59511  
 A/Molecule type: mRNA  
 A/Residues: 302-423 <CHT>  
 A/Cross-references: GB:S79877; NID:g1195576; PIDD:AA53515.1; PID:g1195577  
 R/Seyer, J.M.; Kang, A.H.  
 Biochemistry 17, 3404-3411, 1978  
 A/Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBR pe  
 A/Reference number: A90414; MUID:79000343; PMID:687591  
 A/Accession: A90414  
 A/Molecule type: protein  
 A/Residues: 399-675,'N',677-727 <SEY3>  
 A/Experimental source: liver  
 R/lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
 J. Biol. Chem. 266, 5256-5259, 1991  
 A/Title: G to T transversion at position +5 of a splice donor site causes skipping of th  
 A/Reference number: 155349; MUID:91161621; PMID:1672129  
 A/Accession: 155349  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 537-605 <LEB>  
 A/Cross-references: GB:M59312; NID:g180815; PIDD:AA52041.1; PID:g180816  
 R/Seyer, J.M.; Mainardi, C.; Kang, A.H.  
 Biochemistry 19, 1583-1589, 1980  
 A/Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB5 from cy  
 A/Reference number: A90438; MUID:80198282; PMID:6246925  
 A/Accession: A90438  
 A/Molecule type: protein  
 A/Residues: 728-895,'A',897-964 <SEY4>  
 A/Experimental source: liver  
 R/Cole, W.G.; Chiodo, A.A.; Lamanche, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Chan  
 J. Biol. Chem. 265, 17070-17077, 1990  
 A/Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping an  
 A/Reference number: A38303; MUID:91009133; PMID:2145268  
 A/Accession: A38303  
 A/Molecule type: mRNA  
 A/Residues: 861-1015 <COL>  
 A/Cross-references: GB:U05617; GB:M55603; GB:M59227; NID:g180878; PIDD:AA59383.1; PID:g

A/Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syn  
 R/Mankoo, B.S.; Daigleish, R.  
 Nucleic Acids Res. 16, 2337, 1988  
 A/Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.  
 A/Reference number: S02119; MUID:88189827; PMID:3357782  
 A/Accession: S02119  
 A/Status: translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 950-1018,'Y',1020-1183,'S',1185-1466 <MAN>  
 A/Cross-references: EMBL:X06700; NID:g30053; PIDD:CAA29886.1; PID:g30054  
 R/Seyer, J.M.; Kang, A.H.  
 Biochemistry 20, 2621-2627, 1981  
 A/Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB9 from cy  
 A/Reference number: A90446; MUID:81208139; PMID:7016180  
 A/Accession: A90446  
 A/Molecule type: protein  
 A/Residues: 965-979,'A',981-984,'PS',987,'QN',990-1096,'P',1098-1152,'AT',1155,'S',1157-  
 A/Experimental source: liver  
 R/Loidl, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Mye  
 Nucleic Acids Res. 12, 9383-9394, 1984  
 A/Title: Molecular cloning and carboxyl-propeptide analysis of human type III procollage  
 A/Reference number: A93551; MUID:85087944; PMID:6096827  
 A/Accession: A93551  
 A/Molecule type: mRNA  
 A/Residues: 1065-1155,'P',1157-1466 <LOI>  
 A/Cross-references: EMBL:X01655; EMBL:X01742; NID:g29584; PIDD:CAA25821.1  
 R/Mikulin, M.; Daigleish, R.; Kuve-Beckerman, B.; Renard, S.I.; Tolstoshev, P.; Brant  
 Biochemistry 25, 1408-1413, 1986  
 A/Title: Human type III collagen gene expression is coordinately modulated with the type  
 A/Reference number: 152393; MUID:86187804; PMID:3754462  
 A/Accession: 152393  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1161-1200 <MS>  
 A/Cross-references: GB:M13146; NID:g180415; PIDD:AA52003.1; PID:g180416  
 R/Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.M.; Myers, J.C.  
 Proc. Natl. Acad. Sci. U.S.A. 82, 3385-3389, 1985  
 A/Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm  
 A/Reference number: 159025; MUID:85157600; PMID:3858826  
 A/Accession: 179359  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 1165-1196 <EMA>  
 A/Cross-references: GB:M1134; NID:g180417; PIDD:AA52004.1; PID:g180418  
 R/Chu, M.L.; Weil, D.; de Weert, W.; Bernard, M.; Sipola, M.; Ramirez, F.  
 J. Biol. Chem. 260, 4357-4363, 1985  
 A/Title: Isolation of cDNA and genomic clones encoding human pro-alpha1(III) collagen. P  
 A/Reference number: A92516; MUID:85157600; PMID:2579949  
 A/Accession: A92516  
 A/Molecule type: DNA  
 A/Residues: 1176-1240,'V',1242-1356,'P',1358-1466 <CHU>  
 A/Cross-references: GB:M10615; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB  
 A/Experimental source: liver  
 A/Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given f  
 action  
 C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit (C  
 3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently O  
 C/Genes: GDB:COL3A1  
 A/Gene: GDB:COL3A1  
 A/Cross-references: GDB:118729; OMIM:120180  
 A/Map position: 2q31-2q31  
 A/Intons: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1275/1; 1337/3; 1418/3  
 A/Note: the list of introns is incomplete; defects in this gene can result in Ehlers-Dan  
 C/Complex: type III collagen is a homotrimer of monomers initially linked by disulfide b  
 er of their length, is formed with desmosine cross-links made from lysine and allysine r  
 C/Function:  
 A/Description: structural component of extracellular fibrous polymer that maintains int  
 C/Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;  
 C/Keywords: coiled coil; Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hyd  
 F:1-23/Domain: signal sequence #status predicted <SIG>  
 F:24-153/Domain: amino-terminal propeptide #status predicted <PRO>  
 F:1-91/Domain: von Willebrand factor type C repeat homology <WVC>  
 F:154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>



```
F:154-167/Region: amino-terminal nonhelical telopeptide
F:168-1196/Region: helical
F:1091-1093/Region: cell attachment (R-G-D) motif
F:1197-1221/Region: carboxyl-terminal nonhelical telopeptide
F:1222-1466/Domain: fibrillar collagen propeptide #status predicted <CPR>
F:1238-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>
F:242/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:153-154/Cleavage site: Pro-Gln (procollagen N-endopeptidase) #status predicted
F:154/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F:161-1212/Modified site: allysine (Lys) #status predicted
F:263-284,860,977,1106/Modified site: 5-hydroxylysine (Lys) #status experimental
F:263/Binding site: carboxydrate (Lys) (covalent) #status experimental
F:584-1099/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental
F:948-949/Cleavage site: Gly-Ile (collagenase) #status experimental
F:1106/Binding site: carboxydrate (Lys) (covalent) #status predicted

Query Match          9.3%; Score 121; DB 1; Length 1466;
Best Local Similarity 40.8%; Pred. No. 0.012;
Matches 31; Conservative 7; Mismatches 30; Indels 8; Gaps 3;

QY 54 NMGCLGPGAVGPRDGSFGANGIPGTGPIGRDGFKEKGECLRESF--EESWTPYKQ 110
DB 100 NCGCPGCPKGDPPGPIGPNNDGPIFGQSPSPSPGPIG--ESCPTGPOVNSPQYDS 157
QY 111 CSWSINYGIDLGKIA 126
DB 158 YDYKS--GVAVGGLA 170

RESULT 10
collagen alpha 2(IV) chain precursor - pig roundworm
C:Species: Ascaris suum (pig roundworm)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S16366
R:Petit, J.; Kingston, I.B.
J. Biol. Chem. 266, 16149-16156, 1991
A:Title: The complete primary structure of a nematode alpha-2(IV) collagen and the parti
A:Reference number: S16366; MUID:91340768; PMID:11714907
A:Accession: S16366
A:Molecule type: mRNA
A:Residues: 1-1763 <JB1>
A:Cross-references: UNIPROT:P27393; GB:M67507; NID:g159648; PIDN:AAA18014.1; PID:g159648
C:Genetics:
A:Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;
C:Superfamily: collagen alpha 1(IV) chain
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulfid
F:1-26/Domain: signal sequence #status predicted <SIG>
F:27-1763/Product: collagen alpha 2(IV) chain #status predicted <NHL>
F:27-42/Domain: non-collagenous NH1 #status predicted <NHL>
F:43-1529/Domain: collagenous #status predicted <COL>
F:197-199/Region: cell attachment (R-G-D) motif
F:1530-1763/Domain: carboxyl-terminal nonhelical, NCI #status predicted <NC1>
F:1530-1638/Domain: repeat NCI #status predicted <NC11>
F:1639-1763/Domain: repeat NCI #status predicted <NC12>
F:31-34,39,41,536,539/Disulfide bonds: interchain #status predicted
F:126/Binding site: carboxydrate (Asn) (covalent) #status predicted
F:1593-1599,1702-1709/Disulfide bonds: #status predicted

Query Match          9.3%; Score 121; DB 2; Length 1763;
Best Local Similarity 62.9%; Pred. No. 0.015;
Matches 22; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 60 GPAGVPRDGSFGANGIPGTGPIGRDGFKEKGE 94
DB 162 GYAGAPGKGDADGAPGPIGPPGPRDGFKEKGD 196

RESULT 11
S29328
complement subcomponent C1q chain C - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004

C:Accession: S29328
R:Petty, F.; Reid, K.B.M.; Loos, M.
Eur. J. Biochem. 209, 129-134, 1992
A:Title: Isolation, sequence analysis and characterization of cDNA clones coding for th
e c1q domain.
A:Reference number: S29328; MUID:93011118; PMID:1396691
A:Accession: S29328
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-246 <PET>
A:Cross-references: UNIPROT:Q02105; EMBL:X66295; NID:g50228; PIDN:CAA46993.1; PID:g5022
C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal ho
F:122-245/Domain: complement C1q carboxyl-terminal homology <C1Q>

Query Match          9.1%; Score 119.5; DB 2; Length 246;
Best Local Similarity 36.6%; Pred. No. 0.0023;
Matches 34; Conservative 6; Mismatches 32; Indels 21; Gaps 3;

QY 5 GRASPQLLGFVLVLLQLSAPSSASENPKVKALIRQEVVDLYNMGCLQGPGAV 64
DB 4 GRSCDQ--GGLCLLFLPLALPLRSQAS-----AGCTGIRGMPMPGA 45
QY 65 PGRDGS--PGANGIPGTGPIGRDGFKEKGE 94
DB 46 PGKDGHDLGQPGKPGPIGPAVPTGQPGKQKGE 78

RESULT 12
T29731
hypothetical protein F41F3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29731
R:Padcock, K.; Le, T.T.; Gattung, S.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid F41F3.
A:Reference number: Z20674
A:Accession: T29731
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-310 <BAD>
A:Cross-references: UNIPROT:Q20282; EMBL:U55366; PIDN:AAA97982.1; GSPDB:GN00023; CESP:F
C:Experimental source: strain Bristol N2; clone F41F3
C:Genetics:
A:Gene: CESP:F41F3.4
A:Map position: 5
A:Introns: 27/3; 68/3

Query Match          9.1%; Score 119.5; DB 2; Length 310;
Best Local Similarity 40.7%; Pred. No. 0.003;
Matches 24; Conservative 4; Mismatches 10; Indels 21; Gaps 1;

QY 57 CLOGPAGVPRDGSFGANGIPGTGPI-----PGRDGFKEKGE 94
DB 158 CFPGAPGEGHGDGAPGAPGPGVGGSGAPGAPGKAPGPGQPGRDGPGQAGQ 216

RESULT 13
A61396
collagen alpha 1(II) chain - golden hamster (fragments)
C:Species: Mesocricetus auratus (golden hamster)
C:Date: 09-Sep-1994 #sequence_revision 09-Sep-1994 #text_change 09-Jul-2004
C:Accession: A61396
R:Cizdziel, P.E.; Hosi, J.; Montgomery, J.C.; Wiseman, R.W.; Barrett, J.C.
Mol. Cell Biol. 4, 14-24, 1991
A:Title: Loss of a tumor suppressor gene function is correlated with downregulation of
A:Reference number: A61396; MUID:91182265; PMID:2009131
A:Accession: A61396
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-323 <CTZ>
A:Cross-references: UNIPROT:Q7M099
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology;
```

F:120-323/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
Query Match 9.1%; Score 119.5; DB 2; Length 323;  
Best Local Similarity 20.6%; Pred. No. 0.0031;  
Matches 57; Conservative 20; Mismatches 47; Indels 153; Gaps 10;  
QY 58 LOGPAGVPRGDSRGAN-----GIPGPR----- 80  
DB 9 LGGPAGPGRGSGFGADGPPRGDGAAGVCKDRGETALGAPGAPGPPAGPAGATGKGG 68  
QY 81 -----GIPGRDGFKEGKE-----CLRESPE 101  
DB 69 DRGSGAQQPMGSPGAPGARIGPGPGRGDKGSEGPSSLRHQDVVDATLKSLLNQGE 128  
QY 102 ESKTP-----NYQC--SWSLNYGIDLGKIAECTTKRS----- 135  
DB 129 SIRSPGSGCKNPARTCODLKLCHPEKWSGDYIDPNQ--GCTLDARKAKAFQNMETGTCVY 186  
QY 136 -NSA-----LRLVLF---SGSLRLKCRNACQGRWYTFNGAGC 168  
DB 187 PNPASVPKKNWSSKKEKKHIVQMTFLRLSTGSONITTHCKN----- 231  
QY 169 SGPLPEALIVLDGSPELNSTINIRHTSSVEGLCEG 205  
DB 232 -----STAVYDEAAGNLKKALLIQSNDVEKRAEG 261  
RESULT 14  
CGHUB3  
collagen alpha 3(IV) chain precursor, long splice form - human  
N:Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form  
C:Species: Homo sapiens (man)  
C:Date: 28-Oct-1994 #sequence revision 03-Oct-1995 #text\_change 09-Jul-2004  
A:Accession: A54763; A43928; A44043; A45971; A59786  
R:Maruyama, M.; Leinonen, A.; Mochizuki, T.; Tsygvaason, K.; Reeders, S.T.  
J. Biol. Chem. 269, 23013-23017, 1994  
A:Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpression  
A:Reference number: A54763; MUID:94364994; PMID:8083201  
A:Accession: A54763  
A:Molecule type: mRNA  
A:Residues: 1-1670 <MAR>  
A:Cross-references: UNIPROT:Q01955; GB:X80031; NID:g577563; PID:g577564  
A:Experimental source: Kidney  
R:Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.  
J. Clin. Invest. 89, 593-601, 1992  
A:Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha  
A:Reference number: A43928; MUID:92147878; PMID:1737849  
A:Accession: A43928  
A:Molecule type: mRNA  
A:Residues: 1331-1524, 'I', 1526-1670 <TUR>  
A:Cross-references: GB:M81379  
A:Experimental source: Kidney  
R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 267, 19780-19784, 1992  
A:Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture  
A:Reference number: A44043; MUID:93015826; PMID:1400291  
A:Accession: A44043  
A:Molecule type: DNA; mRNA  
A:Residues: 1386-1670 <QUI>  
A:Cross-references: GB:M92993; NID:g177895; PIDN:AAA21610.1; PID:g177896  
A:Note: sequence extracted from NCBI backbone (NCBI:P:115597)  
R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 269, 17358, 1994  
A:Reference number: A44738; MUID:94274734; PMID:8006044  
A:Content: annotation; extratum; correction to intronic sequence in A44043  
R:Bernal, D.; Quinones, S.; Saus, J.  
J. Biol. Chem. 268, 12090-12094, 1993  
A:Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.  
A:Reference number: A45971; MUID:93260184; PMID:8505332  
A:Accession: A45971  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA

A:Residues: 1427-1444 <BBR>  
A:Note: sequence extracted from NCBI backbone (NCBI:133363); sequence incorrectly ident  
R:Morrisson, K.B.; Maruyama, M.; Yang-Feng, T.L.; Reeders, S.T.  
Am. J. Hum. Genet. 49, 545-554, 1991  
A:Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of  
A:Reference number: A39786; MUID:91353570; PMID:1882840  
A:Accession: A39786  
A:Molecule type: mRNA  
A:Residues: 1433-1593, 'A', 1595-1670 <MOR>  
A:Cross-references: GB:S55790; NID:g234418; PIDN:AB19637.1; PID:g234419  
A:Comment: Prolines and lysines at the third position of the tripeptide repeating unit (ed and subsequently O-glycosylated).  
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope  
C:Genetics:  
A:Gene: GDB:COL4A3  
A:Cross-references: GDB:128351; OMIM:120070  
A:Map position: 2q36-q37  
A:Positions: 1385/1; 1418/1; 1488/1; 1547/2; 1585/3; 1643/2 #status incomplete  
A:Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3 (m  
er trimmer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-  
C:Function:  
A:Description: minor structural component of extracellular basement membrane in kidney  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracel  
F:1-28/Domain: signal sequence #status predicted <SIG>  
F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAR>  
F:29-42/Domain: amino-terminal nonhelical, NH1 <NH1>  
F:1433-1438/Region: interrupted helical  
F:791-793/Region: cell attachment (R-G-D) motif  
F:996-998/Region: cell attachment (R-G-D) motif  
F:1154-1156/Region: cell attachment (R-G-D) motif  
F:1306-1308/Region: cell attachment (R-G-D) motif  
F:1345-1347/Region: cell attachment (R-G-D) motif  
F:1432-1434/Region: cell attachment (R-G-D) motif  
F:1433-1670/Domain: carboxyl-terminal nonhelical, NCI <NCI>  
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>  
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CT2>  
F:31,33,39,41,125,422,476,479,662,722,809,1387/Disulfide bonds: interchain #status predi  
F:553/Binding site: carbonyl-ate (asn) (covalent) #status predicted  
F:1460-1548, 1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted  
F:1505-1511, 1616-1622/Disulfide bonds: #status predicted  
F:1570-1662, 1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted  
Query Match 9.1%; Score 119.5; DB 1; Length 1670;  
Best Local Similarity 57.1%; Pred. No. 0.02;  
Matches 24; Conservative 3; Mismatches 12; Indels 3; Gaps 1;  
QY 59 GGPAGVPRGDSRGANIGRTGIGPRD---GFKGKCECLR 97  
DB 504 QGAGLKGSGSGSPGNTGLGPFPGAGQDPGLKKEKERTLQ 545  
RESULT 15  
A54849  
collagen alpha 1(VII) chain precursor - human  
N:Alternate names: procollagen alpha 1(VII) chain  
C:Species: Homo sapiens (man)  
C:Date: 04-Nov-1994 #sequence revision 04-Nov-1994 #text\_change 09-Jul-2004  
A:Accession: A54849; P00844; S16316; I56328; A30296; I84686  
R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.  
J. Biol. Chem. 269, 20256-20262, 1994  
A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha (VII  
A:Reference number: A54849; MUID:94327588; PMID:8051117  
A:Accession: A54849  
A:Molecule type: mRNA  
A:Status: not compared with conceptual translation  
A:Status: not compared with conceptual translation  
A:Reference number: UNIPROT:Q02388; GB:I02870; NID:9987124; PIDN:AAA5438.1; PID:g987125  
R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Imamura, S.  
Biochem. Biophys. Res. Commun. 183, 958-963, 1992  
A:Title: Molecular cloning and characterization of type VII collagen cDNA.

A/Reference number: PH0844; MUID:92231902; PMID:1567409  
 A/Accession: PH0844  
 A/Molecule type: mRNA  
 A/Residues: 'EPR',340-475,'RALSTASHSTICWRAETWHPNCRSHWTTRACEDCPNPASHRAARAG',524-528,'C',  
 A/Cross-references: DDBJ:DJ11152; DDBJ:DJ13694; NID:9455698; PIDN:BA02853.1; PID:9453699  
 A/Experimental source: Keratinocyte  
 A/Note: the authors translated the codon ACC for residues 394 and 397 as Tyr  
 R/Parete, M.G.; Chung, L.C.; Ryttaenen, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat  
 Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991  
 A/Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.  
 A/Reference number: S16316; MUID:91334380; PMID:1871109  
 A/Accession: S16316  
 A/Molecule type: mRNA  
 A/Residues: 815-892,'E',894-1439,<PAR>  
 A/Cross-references: GB:M65158; GB:S49017; NID:9180914; PIDN:AAA96439.1; PID:9180915  
 A/Experimental source: Keratinocyte  
 R/Gemmon, W.R.; Abernethy, M.L.; Padilla, K.M.; Pristayn, P.S.; Cook, M.E.; Wright, J.  
 J. Invest. Dermatol. 99, 691-696, 1992  
 A/Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot  
 A/Reference number: I56328; MUID:93107742; PMID:1469284  
 A/Accession: I56328  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 'EPR',372-517,'DV',520-540,'W',542-1255,<RES>  
 A/Cross-references: GB:S51236; NID:9262308; PIDN:AB24637.1; PID:9262309  
 R/Salter, J.L.; Eissen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgess, R.E.  
 J. Biol. Chem. 264, 3822-3826, 1989  
 A/Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagena  
 A/Reference number: A30296; MUID:89139437; PMID:2537292  
 A/Accession: A30296  
 A/Molecule type: protein  
 A/Residues: 'A',1240-1246,'G',1248-1250,'XE',1253-1255,'Q',1257,'E',2032,'C',2034-2041,'  
 A/Note: two reported peptides cannot be reliably located  
 R/Greenspan, D.S.  
 Hum. Mol. Genet. 2, 273-278, 1993  
 A/Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous  
 A/Reference number: I48103; MUID:93271985; PMID:8499916  
 A/Accession: I48103  
 A/Status: preliminary; translated from GB/EMBL/DBJ  
 A/Molecule type: mRNA  
 A/Residues: 2395-2871,'S',2873-2944,<RE2>  
 A/Cross-references: GB:L06862; NID:9388713; PIDN:AAA89196.1; PID:9388714  
 R/Christiano, A.M.; Ryttaenen, M.; Uitto, J.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994  
 A/Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser sub  
 A/Reference number: A55255; MUID:94224777; PMID:8170945  
 A/Contents: annotation  
 C/Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 ed and subsequently O-glycosylated.  
 C/Genetics:  
 A/Genes: GDB:COL7A1; EBR1; EBD1; EB  
 A/Cross-references: GDB:128750; OMIM:120120  
 A/Map position: 3p21.3-3p21.3  
 A/Note: defects in this gene can result in dominant and recessive dystrophic epidermolys  
 A/Note: there are 118 introns  
 C/Complex: type VII collagen is probably a homotrimer  
 C/Function:  
 A/Description: Structural component of extracellular polymer associated with anchoring  
 C/Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyproli  
 F/1-16/Domain: signal sequence #status predicted <SIG>  
 F/17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>  
 F/17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>  
 F/36-201/Domain: von Willebrand factor type A repeat homology <VWA1>  
 F/327-413/Domain: fibronectin type III repeat homology <FN1>  
 F/414-502/Domain: fibronectin type III repeat homology <FN2>  
 F/508-593/Domain: fibronectin type III repeat homology <FN3>  
 F/598-683/Domain: fibronectin type III repeat homology <FN4>  
 F/686-771/Domain: fibronectin type III repeat homology <FN5>  
 F/776-862/Domain: fibronectin type III repeat homology <FN6>  
 F/864-952/Domain: fibronectin type III repeat homology <FN7>  
 F/954-1045/Domain: fibronectin type III repeat homology <FN8>  
 F/1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>

F/1170-1172/Region: cell attachment (R-G-D) motif  
 F/1189-1253/Region: cysteine/proline-rich  
 F/1254-2783/Region: interrupted helical  
 F/1334-1336/Region: cell attachment (R-G-D) motif  
 F/2008-2010/Region: cell attachment (R-G-D) motif  
 F/2553-2555/Region: cell attachment (R-G-D) motif  
 F/2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>  
 F/2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
 F/337,786,1109/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F/2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status expe  
 F/2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental  
 F/2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental  
 F/2634,2802,2804/Disulfide bonds: interchain #status predicted

Query Match 9.1%; Score 119.5; DB 2; Length 2944;  
 Best Local Similarity 60.5%; Pred. No. 0.037;  
 Matches 23; Conservative 3; Mismatches 9; Indels 3; Gaps 1;  
 QY 60 GPAGVPCR---DQSPGANGIPGTPGIPGSDGKGEKE 94  
 DB 1824 GPGPLGKPGEDGKPGELNGKNGSPGDPGSDGKGEKXD 1861

Search completed: August 4, 2005, 04:22:28  
 Job time : 20.5359 secs

**This Page Blank (uspto)**



CC hypertrophic scar formation. REMODELIN is useful in the development of  
CC bone during mammalian embryogenesis, proliferation and/or migration.  
CC cellular signaling, adventitial fibrosis, negative remodeling and arterial  
CC restenosis, smooth muscle cell proliferation, and arterial remodeling and  
CC useful in gene therapy. REMODELIN is useful for elucidating the function  
CC of REMODELIN molecules in a cell, to identify a compound that affects  
CC REMODELIN expression and/or TGF-beta signaling, as a potential  
CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
CC formation, to promote bone fracture healing, and to increase or decrease  
CC expression of REMODELIN in mammal. It is also useful for producing  
CC recombinant cell and transgenic non-human mammals which are useful tools  
CC for the study of REMODELIN action, for identifying novel diagnostic and  
CC therapeutics for treatment, and for elucidating the cellular roles of  
CC REMODELIN. The present sequence represents a rat REMODELIN polypeptide  
XX  
XX  
SQ Sequence 245 AA;

Query Match 100.0%; Score 1307; DB 5; Length 245;  
Best Local Similarity 100.0%; Pred. No. 3,7e-123;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPQLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
DB 1 MHPOGRAASPQLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
QY 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNPKQCSWSSLYNGI 120  
DB 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNPKQCSWSSLYNGI 120  
QY 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWFTFNAGCSGPIPIEAIITL 180  
DB 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWFTFNAGCSGPIPIEAIITL 180  
QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRILII 240  
DB 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRILII 240  
QY 241 EELPK 245  
DB 241 EELPK 245

RESULT 2  
ABG75757  
ID ABG75757 standard; protein; 245 AA.

AC ABG75757;  
XX  
DT 28-APR-2003 (first entry)

DE Rat REMODELIN protein.

XX Rat; adventitia-inducible; REMODELIN; REMODEL;  
KM adventitia induced bone expressed molecule; AIBE; TGF-beta;  
KM transforming growth factor beta; adventitia; vascular remodeling;  
KM restenosis; vascular injury; antisense therapy; TGF-beta signaling;  
KM TGF-beta receptor type II; arterial remodeling; bone formation;  
KM cartilage formation; osteogenesis imperfecta; Bethlem myopathy;  
KM dystrophic epidermolysis bullosa; negative remodeling; wound healing;  
KM arterial stenosis; fibrosis; calcification; transplant;  
KM heart valve transplant; osteopathic; antiteriosclerotic; vulnerary.

OS Rattus sp.

XX US2002161211-A1.

PD 31-OCT-2002.

XX 19-OCT-2001; 2001US-00045992.

PR 19-OCT-2000; 2000US-0062081.

XX

PA (LIND/) LINDNER V.  
PA (FRIE/) FRIESEL R E.  
XX  
XX Lindner V, Frieel RE;  
XX  
XX MPI: 2003-238238/23.  
DR N-PSDB; ABX11340.

PT New isolated REMODELIN nucleic acid and polypeptide, useful for mediating  
PT arterial remodeling, formation of bone and cartilage, and the diagnosis  
PT and treatment of disorders associated with aberrant expression of  
PT REMODELIN.

PS Claim 3; Fig 4B; 81pp; English.

XX The invention discloses an isolated nucleic acid encoding a mammalian  
CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL  
CC and/or adventitia induced bone expressed molecule, AIBE). REMODELIN  
CC expression was induced by transforming growth factor beta (TGF-beta),  
CC which is important because proliferative events occurring in the  
CC adventitia contribute to vascular remodeling and restenosis in response  
CC to vascular injury and TGF-beta has been shown to be a factor involved in  
CC this. Also disclosed is an antibody raised against REMODELIN and methods  
CC for treating a disease mediated by abnormal expression of a REMODELIN in  
CC a human, for identifying a compound that affects or reduces expression of  
CC REMODELIN in a cell (e.g. antisense therapy), for identifying a compound  
CC that affects TGF-beta signaling and for increasing or reducing REMODELIN  
CC expression in a mammal, comprising administering a REMODELIN expression  
CC increasing or reducing TGF-beta to the mammal, thereby increasing  
CC REMODELIN expression or inhibiting signaling via the TGF-beta receptor  
CC type II and reducing expression of REMODELIN in the mammal. The method  
CC and compositions of the present invention are useful for mediating  
CC arterial remodeling, formation of bone and cartilage and the diagnosis  
CC and treatment of disorders associated with aberrant expression of  
CC REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis  
CC bullosa, Bethlem myopathy, negative remodeling, wound healing, arterial  
CC stenosis, vessel injury, fibrosis and calcification of a transplant,  
CC preferably a heart valve transplant. The sequence presented is the rat  
CC REMODELIN protein  
XX

SQ Sequence 245 AA;

Query Match 100.0%; Score 1307; DB 6; Length 245;  
Best Local Similarity 100.0%; Pred. No. 3,7e-123;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPQLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
DB 1 MHPOGRAASPQLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60  
QY 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNPKQCSWSSLYNGI 120  
DB 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGFKEGECLEARSFESWTPNPKQCSWSSLYNGI 120  
QY 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWFTFNAGCSGPIPIEAIITL 180  
DB 121 DGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCORWFTFNAGCSGPIPIEAIITL 180  
QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRILII 240  
DB 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMNSVSRILII 240  
QY 241 EELPK 245  
DB 241 EELPK 245

RESULT 3

ABB80979  
ID ABB80979 standard; protein; 277 AA.

AC ABB80979;

XX

DT	21-OCT-2002	(first entry)
XX		
DE	Long form of rat REMODELIN polypeptide.	
XX		
KV	REMODELIN; vulnery; vasotropic; cytostatic; osteopathic; collagen,	
KW	gene therapy; bone; rat.	
XX		
OS	Rattus sp.	
XX		
PN	WO200242487-A2.	
XX		
PD	30-MAY-2002.	
XX		
PF	19-OCT-2001; 2001WO-US050940.	
XX		
PR	19-OCT-2000; 2000US-00692081.	
XX		
PA	(MAIN-) MAINE MEDICAL CENT RES INST.	
XX		
PI	Linder V, Friesel R;	
DR	WPI; 2002-590472/63.	
DR	N-PSDB; ABR86480.	
XX		
PT	New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,	
PT	negative remodeling, fibrosis, collagen disease, and bone disease in a	
XX	mammal.	
PS	Claim 5; Fig 4C; 187P; English.	
XX		
CC	The invention relates to mammalian REMODELIN polypeptides and encoding	
CC	polynucleotides. REMODELIN is used to diagnose arterial restenosis,	
CC	negative remodeling or fibrosis, bone disease such as osteogenesis	
CC	imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis	
CC	bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are	
CC	used to affect cellular gene expression in a mammal, where the cellular	
CC	gene is from transforming growth factor (TGF)-beta1, collagen I(1a1p1a1,	
CC	osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein	
CC	4. The expression of osteopontin is dependent on Cbfa1. Compositions	
CC	comprising antisense REMODELIN sequences are useful for treating diseases	
CC	mediated by abnormal expression of a REMODELIN molecule in a human such	
CC	as impaired wound healing, fibrosis of an organ, ectopic ossification, or	
CC	hypertrophic scar formation. REMODELIN is useful in the development of	
CC	bone during mammalian embryogenesis, proliferation and/or migration,	
CC	cellular signaling, adventitial fibrosis, negative remodeling and arterial	
CC	restenosis, smooth muscle cell proliferation, and arterial remodeling and	
CC	useful in gene therapy. REMODELIN is useful for elucidating the function	
CC	of REMODELIN molecules in a cell, to identify a compound that affects	
CC	REMODELIN expression and/or TGF-beta signaling, as a potential	
CC	therapeutic drug candidate for arterial restenosis, anti-cancer therapy,	
CC	to promote or inhibit wound healing, to inhibit scar tissue or keloid	
CC	formation, to promote bone fracture healing, and to increase or decrease	
CC	expression of REMODELIN in mammal. It is also useful for producing	
CC	recombinant cell and transgenic non-human mammals which are useful tools	
CC	for the study of REMODELIN action, for identifying novel diagnostics and	
CC	therapeutics for treatment, and for elucidating the cellular roles of	
CC	REMODELIN. The present sequence represents the long form of rat REMODELIN	
CC	polypeptide	
XX		
SQ	Sequence 277 AA;	
Query Match	100.0%; Score 1307; DB 5; Length 277;	
Best Local Similarity	100.0%; Pred. No. 4.4e-12;	
Matches 245; Conservative	0; Mismatches 0; Indels 0; Gaps 0	
QY	1 MHPOGRAASPOLLLIGLFTVLLILLLOLSAPSSASSENPKVOKALLIROREVVDYNGMCLG 60	
Db	33 MHPQGRASPOLLLIGLFTVLLILLLOLSAPSSASSENPKVOKALLIROREVVDYNGMCLG 92	
QY	61 PAVGVRGSGSPRANGIPGPIPGSDGFGKGGCECRSEFSSWYNYQCSNLSNYGI 120	
Db	93 PAVGVRGSGSPRANGIPGPIPGSDGFGKGGCECRSEFSSWYNYQCSNLSNYGI 152	

Oy	121	DGKIAECTFTYKRRNSALRVLFSSGLRLKCRNACCORWFTFNGAECSSGPLPEALIIYL	180
Oy	121	DGKIAECTFTYKRRNSALRVLFSSGLRLKCRNACCORWFTFNGAECSSGPLPEALIIYL	180
Db	153	DGKIAECTFTYKRRNSALRVLFSSGLRLKCRNACCORWFTFNGAECSSGPLPEALIIYL	212
Oy	181	DQSSPELNTNTINHRSSVEGICEGTIGGLYDVVAIHWGTCSDPYKGDASTGMSVSRITII	240
Db	213	DQSSPELNTNTINHRSSVEGICEGTIGGLYDVVAIHWGTCSDPYKGDASTGMSVSRITII	272
Oy	241	EEELPK 245	
Db	273	EEELPK 277	
RESULT 4			
ABG75759			
ID	ABG75759	standard; protein; 277 AA.	
AC			
XX	ABG75759;		
XX			
DT	28-APR-2003	(first entry)	
XX			
DE	Rat REMODELIN long form (REMODELINL) protein.		
XX			
KM	Rat; adventitia-inducible; REMODELIN; rREMODELINL; REMODEL;		
KM	adventitia induced bone expressed molecule; AIBE; TGF-beta;		
KM	transforming growth factor beta; adventitia; vascular remodelling;		
KM	restenosis; vascular injury; antisense therapy; TGF-beta signalling;		
KM	TGF-beta receptor type II; arterial remodelling; bone formation;		
KM	cartilage formation; osteogenesis imperfecta; Bethlem myopathy;		
KM	dystrophic epidermolysis bullosa; negative remodelling; wound healing;		
KM	arterial stenosis; fibrosis; calcification; transplant;		
KM	heart valve transplant; osteopathic; antiarteriosclerotic; vulnerary.		
OS			
XX	Rattus sp.		
XX			
FN	US2002161211-A1.		
XX			
PD	31-OCT-2002.		
XX			
PE	19-OCT-2001; 2001US-00045992.		
XX			
FR	19-OCT-2000; 2000US-00692081.		
XX			
PA	(LIND/) LINDNER V.		
PA	(FRIE/) FRIESEL R E.		
FI	Lindner V, Friesel RE;		
XX			
DR	WPI; 2003-238238/23.		
XX			
DR	N-PSDB; ABX11340.		
XX			
PS			
XX	Claim 3; Fig 4C; 81pp; English.		
CC	The invention discloses an isolated nucleic acid encoding a mammalian		
CC	adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL		
CC	and/or adventitia induced bone expressed molecule, AIBE). REMODELIN		
CC	expression was induced by transforming growth factor beta (TGF-beta),		
CC	which is important because proliferative events occurring in the		
CC	adventitia contribute to vascular remodelling and restenosis in response		
CC	to vascular injury and TGF-beta has been shown to be a factor involved in		
CC	this. Also disclosed is an antibody raised against REMODELIN and methods		
CC	for treating a disease mediated by abnormal expression of a REMODELIN in		
CC	a human, for identifying a compound that affects or reduces expression of		
CC	REMODELIN in a cell (e.g. antisense therapy), for identifying a compound		
CC	that affects TGF-beta signalling and for increasing or reducing REMODELIN		
CC	expression in a mammal, comprising administering a REMODELIN expression		
CC	increasing or reducing TGF-beta to the mammal, thereby increasing		
CC	REMODELIN expression or inhibiting signalling via the TGF-beta receptor		

CC type II and reducing expression of REMODELIN in the mammal. The methods  
 CC and compositions of the present invention are useful for mediating  
 CC arterial remodeling, formation of bone and cartilage and the diagnosis  
 CC and treatment of disorders associated with aberrant expression of  
 CC REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis  
 CC bullousa, Bethlem myopathy, negative remodeling, wound healing, arterial  
 CC stenosis, vessel injury, fibrosis and calcification of a transplant,  
 CC preferably a heart valve transplant. The sequence presented is the long  
 CC form of the rat REMODELIN protein (rREMODELIN).

XX Sequence 277 AA;

Query Match 100.0%; Score 1307; DB 6; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 4,4e-123; Indels 0; Gaps 0;  
 Matches 245; Conservative 0; Mismatches 0;

QY 1 MHPQGRASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 60  
 |||||  
 DB 33 MHPQGRASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 92  
 QY 61 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKEGECLEBSFEBSWTPTYKQCSWSLNYGI 120  
 |||||  
 DB 93 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKEGECLEBSFEBSWTPTYKQCSWSLNYGI 152  
 QY 121 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 180  
 |||||  
 DB 153 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 212  
 QY 181 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNMVSRIII 240  
 |||||  
 DB 213 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNMVSRIII 272  
 QY 241 EELPK 245  
 |||||  
 DB 273 EELPK 277

RESULT 5

ADO28703 standard; protein; 244 AA.

XX ADO28703;  
 XX  
 DT 12-AUG-2004 (first entry)  
 XX  
 DE HA4, SEQ ID 2.  
 XX  
 XX Osteopathic; Cytostatic; Antiinflammatory; Antiarthritic; Antirheumatic;  
 KW osteoblast- and chondrocyte-specific small secreted peptide; HA4;  
 KW bone degenerative disease; osteoporosis;  
 KW glucocorticoid induced osteoporosis; Paget's disease;  
 KW periodontal disease; tooth loss; bone fracture; rheumatoid arthritis;  
 KW periprosthetic osteolysis; osteogenesis imperfecta;  
 KW metastatic bone disease; hypercalcemia of malignancy.  
 OS  
 XX Unidentified.  
 XX  
 PN MO2004041205-A2.  
 XX  
 PD 21-MAY-2004.  
 XX  
 PF 04-NOV-2003; 2003WO-US035139.  
 XX  
 PR 04-NOV-2002; 2002US-0423690P.  
 XX  
 XX (TEXA ) UNIV TEXAS SYSTEM.  
 PA  
 PI De Crombrughe B, Akiyama H;  
 XX  
 XX WPI; 2004-411407/38.  
 DR  
 DR N-PSDB; ADO28702.  
 XX  
 PT Novel isolated osteoblast- and chondrocyte-specific small secreted

PT peptide referred as HA4 polypeptide, useful for diagnosing or treating  
 PT bone degenerative disease, and for screening drugs that alter bone  
 PT formation.

XX Claim 1; SEQ ID NO 2; 103p; English.

CC The present invention relates to an isolated osteoblast- and chondrocyte-  
 CC specific small secreted peptide named HA4 polypeptide (I; ADO28703) and  
 CC its coding sequence (II; ADO28702). (I) or (II) are useful for diagnosing  
 CC or treating a bone degenerative disease in a subject such as  
 CC osteoporosis, glucocorticoid induced osteoporosis, Paget's disease,  
 CC abnormally increased bone turnover, periodontal disease, tooth loss, bone  
 CC fractures, rheumatoid arthritis, periprosthetic osteolysis, osteogenesis  
 CC imperfecta, metastatic bone disease and hypercalcemia of malignancy. (II)  
 CC is also useful for identifying a subject at risk of or suffering from a  
 CC bone degenerative disease.

XX Sequence 244 AA;

Query Match 98.5%; Score 1287.5; DB 8; Length 244;  
 Best Local Similarity 99.2%; Pred. No. 3,4e-121; Indels 1; Gaps 1;  
 Matches 243; Conservative 1; Mismatches 1;

QY 1 MHPQGRASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 60  
 |||||  
 DB 1 MHPQGRASPOLLGLFLVLLLLQLSAPSSASENPKVKQALIRQREVVDLYNGMCLQG 59  
 QY 61 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKEGECLEBSFEBSWTPTYKQCSWSLNYGI 120  
 |||||  
 DB 60 PAGVPRDQSPGANGIPGTPGIPGRDGFGEKEGECLEBSFEBSWTPTYKQCSWSLNYGI 119  
 QY 121 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 180  
 |||||  
 DB 120 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCQWYFTFNAGCSGPIPIEAIYYL 179  
 QY 181 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNMVSRIII 240  
 |||||  
 DB 180 DQSPPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNMVSRIII 239  
 QY 241 EELPK 245  
 |||||  
 DB 240 EELPK 244

RESULT 6

AAB36667 standard; protein; 243 AA.

XX AAB36667;  
 XX  
 DT 14-MAR-2001 (first entry)  
 XX  
 DE Human secretory protein TGC-628 SEQ ID NO:7.  
 XX  
 XX Human; secretory protein; cancer; immune disease; infectious disease;  
 KW lung function disorder; liver function disorder; antiinflammatory;  
 KW gastrointestinal disorder; cytostatic; haematopoietic; anticoagulant;  
 KW immunomodulatory; hepatotropic; cell proliferation-stimulant;  
 KW cell migratory agent; cell differentiation-inducer.  
 OS  
 XX Homo sapiens.  
 XX  
 PN MO200071581-A1.  
 XX  
 PD 30-NOV-2000.  
 XX  
 PF 19-MAY-2000; 2000WO-JP003221.  
 XX  
 PR 20-MAY-1999; 99JP-00140229.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 PI Itoh Y, Mogi S, Tanaka H, Ohkubo S, Ogi K;



```

XX WPI: 2001-032023/04.
DR N-PSDB; AAC90707.
PT Novel secretory protein and its salt with e.g. anti-cancer, anti-
XX inflammatory and hematopoietic, effects, applicable as drugs in remedies
XX and preventives to treat diseases like cancer and immune diseases.
PS Claim 1, Page 89-90; 122pp; Japanese.
CC AAC90701 to AAC90715 encode the human secretory proteins given in
CC AAB3661 to AAB3675. The proteins can have cytostatic, anti-
CC inflammatory, haematopoietic, anti-coagulant, immunomodulatory and
CC hepatotropic activities, and can be used as cell migratory agents, cell
CC proliferation- stimulants and cell differentiation-inducers. The proteins
CC are useful in the treatment and prevention of diseases such as cancer,
CC lung function disorder, liver function disorder, gastrointestinal
CC disorder and immune diseases. AAC90716 to AAC90755 represent PCR primers
XX which are used in the exemplification of the present invention
SQ Sequence 243 AA;

Query Match          93.7%; Score 1225; DB 4; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.9e-115;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASBNPKYKQALRQREVVDLYNGCLOG 60
DB 1 MRQGPASPOLRRG--LTLTLLLLQLPAPSSASBIPKQKQALRQREVVDLYNGCLOG 58
QY 61 PAGVGRDSSPGANGIPGTGIPGRDGFKEGECLEKESPEESMTPNYKQCSWSLNYGI 120
DB 59 PAQVGRDSSPGANGIPGTGIPGRDGFKEGECLEKESPEESMTPNYKQCSWSLNYGI 118
QY 121 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 180
DB 119 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 178
QY 181 DOGSPELNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 240
DB 179 DOGSPENNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 238
QY 241 EELPK 245
DB 239 EELPK 243

RESULT 7
ABG96340
ID ABG96340 standard; protein; 243 AA.
XX
AC ABG96340;
XX
DT 11-DEC-2002 (first entry)
XX
DE Human ovarian cancer marker M450.
XX
KW Human; ovarian cancer; marker; cancer; familial history; brain disorder;
KW central nervous system disorder; bacterial meningitis; viral meningitis;
KW Alzheimer's disease; Parkinson's disease; cerebral oedema; hydrocephalus;
KW brain herniation; inflammation; encephalitis; testicular disorder;
KW nondercutaneous granulomatous orchitis; connective tissue disorder;
KW heart disorder; ischaemic heart disease; atherosclerosis; neoplasm;
KW histological type; carcinogenic; ovarian cancer marker.
XX
OS Homo sapiens.
XX
PN MO200271928-A2.
XX
PD 19-SEP-2002.
XX
PF 14-MAR-2002; 2002MO-US007826.
XX

```

```

PR 14-MAR-2001; 2001US-0276025P.
PR 14-MAR-2001; 2001US-0276026P.
PR 10-AUG-2001; 2001US-0311732P.
PR 19-SEP-2001; 2001US-0323580P.
PR 26-SEP-2001; 2001US-0324967P.
PR 26-SEP-2001; 2001US-0325102P.
PR 26-SEP-2001; 2001US-0325149P.
XX
XX (MILL-) MILLENNium PHARM INC.
XX
XX Monahan JE, Ganavavaru M, Hoersch S, Kamatkar S, Kovatis SG,
XX Meyers RE, Morrisey MP, Qiantu PU, Sen A, Vieby PO, Mills GB,
XX Bast RC, Lu K, Schmandt RE, Zhao X, Glatz K;
DR WPI: 2002-723277/78.
XX N-PSDB; AB576436.
XX
XX Assessing whether a patient is afflicted with ovarian cancer, useful in
XX assessing the stage or progression of the disease, comprises comparing
XX the expression level of a cancer marker in a sample from a patient and
XX from a non cancer patient.
XX
XX Disclosure; Page 252; 481pp; English.
XX
XX The present invention relates to a new method for assessing whether a
XX patient is afflicted with ovarian cancer. The method involves comparing
XX the expression level of a marker in a patient sample and the normal level
XX of expression of the marker in a control non-ovarian cancer sample, where
XX the marker is selected from 363 cancer markers described in the
XX specification. The method of the invention is useful in diagnosing or
XX characterizing cancer, in detecting the presence of cancer as early as
XX possible, and the recurrence of ovarian cancer. The method may also be of
XX particular use with patients having an enhanced risk of developing
XX ovarian cancer (e.g. patients having a familial history of ovarian
XX cancer). The cancer markers may be used in the management and treatment
XX of e.g. brain and central nervous system disorders (e.g. bacterial and
XX viral meningitis, Alzheimer's disease or Parkinson's disease), brain
XX disorders (e.g. cerebral oedema, hydrocephalus or brain herniations),
XX inflammations (e.g. bacterial or viral meningitis or encephalitis),
XX testicular disorders (e.g. nondercutaneous granulomatous orchitis),
XX connective tissue disorders, or heart disorders (e.g. ischaemic heart
XX disease or atherosclerosis). The compositions and methods may also be
XX used in assessing the histological type of neoplasm associated with
XX ovarian cancer, monitoring the progression of ovarian cancer, determining
XX whether ovarian cancer has metastasized or is likely to metastasize,
XX selecting a composition for inhibiting ovarian cancer, assessing the
XX ovarian carcinogenic potential of a compound, or inhibiting ovarian
XX cancer or at risk of developing ovarian cancer. The present amino acid
XX sequence represents one of the ovarian cancer markers described in the
XX invention
SQ Sequence 243 AA;

Query Match          93.7%; Score 1225; DB 5; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.9e-115;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASBNPKYKQALRQREVVDLYNGCLOG 60
DB 1 MRQGPASPOLRRG--LTLTLLLLQLPAPSSASBIPKQKQALRQREVVDLYNGCLOG 58
QY 61 PAGVGRDSSPGANGIPGTGIPGRDGFKEGECLEKESPEESMTPNYKQCSWSLNYGI 120
DB 59 PAQVGRDSSPGANGIPGTGIPGRDGFKEGECLEKESPEESMTPNYKQCSWSLNYGI 118
QY 121 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 180
DB 119 DLGKIAECTFTKMRNSALRVLFSGSLRLKCRNACCORWYFTNGAECGSPLEIAIYVL 178
QY 181 DOGSPELNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 240
DB 179 DOGSPENNSTINIHRTSSVSGLCBEGIGAGLVDAIWTGTCSDPYKGDASTGMNSVSRIT 238

```

QY 241 EELPK 245  
|||  
Db 239 EELPK 243

RESULT 8  
ID ABJ05554 standard; protein; 243 AA.  
XX  
AC ABJ05554;  
XX  
DT 14-NOV-2002 (first entry)  
XX  
DE Breast cancer-associated protein 19.  
XX  
KW Breast cancer; breast cancer-associated gene sequence; drug development;  
XX pharmacogenetics; biosensor development.  
XX  
OS Unidentified.  
XX  
PN MO200259377-A2.  
XX  
PD 01-AUG-2002.  
XX  
PF 24-JAN-2002; 2002MO-US002242.  
XX  
PR 24-JAN-2001; 2001US-0263965P.  
XX 02-FEB-2001; 2001US-0265928P.  
XX 09-APR-2001; 2001US-00829472.  
XX 09-APR-2001; 2001US-0282698P.  
XX 04-MAY-2001; 2001US-0288590P.  
XX 29-MAY-2001; 2001US-0294443P.  
XX  
PA (EOSB-) EOS BIOTECHNOLOGY INC.  
XX  
PI Mack DH, Gish KC, Afar D;  
XX  
DR WPI; 2002-583738/62.  
XX N-PSDB; ABT07711.  
XX  
PT Detecting a breast cancer-associated transcript in a patient's cell,  
XX useful for diagnosing breast cancer, comprises contacting a biological  
XX sample with a polynucleotide that selectively hybridizes with breast  
XX cancer nucleic acids.

PS Disclosure; Page 364; 414pp; English.

XX  
XX The invention comprises a method of detecting a breast cancer-associated  
XX transcript in a cell from a patient. The method of the invention involves  
XX contacting a biological sample from the patient with a nucleotide that  
XX hybridizes to one of the 69 breast cancer-associated gene sequences shown  
XX in the specification. The method of the invention is useful in the  
XX diagnosis or prognosis of breast cancer, and for detecting genes that are  
XX up or down-regulated in breast cancer cells. Genes identified by the  
XX method of the invention can be used in diagnostic purposes and also as  
XX targets for screening for therapeutic compounds that modulate breast  
XX cancer (e.g. hormones or antibodies). Identification of genes that are  
XX over or under expressed in breast cancer can additionally provide high-  
XX resolution, high-sensitivity datasets which can be used in the areas of  
XX diagnostics, therapeutics, drug development, pharmacogenetics, protein  
XX structure and biosensor development. Amino acid sequences ABJ05536 -  
XX ABJ05604 represent the proteins encoded by the 69 breast cancer-  
XX associated genes of the invention

SO Sequence 243 AA;

Query Match 93.7%; Score 1225; DB 5; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHQGRAPASPOLLLGLFVLLVLLLLQLSAPSSASNPVVKKALIRQEVVDLYNMCLOG 60  
|||  
Db 1 MRQGPAPASPOLRGG--LTLTLLLLQLPAPSSASEIRPKGKQKQLRQEVVDLYNMCLOG 58

QY 61 PAGVPGRDSPGANGIPGPIGRDGFKGXGECLARESPESWTPNYKQCSWSSLNYGI 120  
|||||  
Db 59 PAGVPGRDSPGANGIPGPIGRDGFKGXGECLARESPESWTPNYKQCSWSSLNYGI 118

QY 121 DLGKIACEFTTMRNSALRVLFSGSLRLKCNACCORWFTFNAGCSGPLPIEAIYYL 180  
|||||  
Db 119 DLGKIACEFTTMRNSALRVLFSGSLRLKCNACCORWFTFNAGCSGPLPIEAIYYL 178

QY 181 DQSGPELNSTINIHRTSSVEGLCEGIGAGLYDVAVIWTGCSDFPRGDASTGNYSIRIIT 240  
|||||  
Db 179 DQSGPELNSTINIHRTSSVEGLCEGIGAGLYDVAVIWTGCSDFPRGDASTGNYSIRIIT 238

QY 241 EELPK 245  
|||  
Db 239 EELPK 243

RESULT 9  
ID ABB80978 standard; protein; 243 AA.  
XX  
AC ABB80978;  
XX  
DT 21-OCT-2002 (first entry)  
XX  
DE Human REMODELIN polypeptide.  
XX  
XX REMODELIN; vulnereary; vasotrophic; cytostatic; osteopathic; collagen;  
XX gene therapy; bone; human.  
XX  
OS Homo sapiens.  
XX  
PN MO200242487-A2.  
XX  
PD 30-MAY-2002.  
XX  
PF 19-OCT-2001; 2001MO-US050940.  
XX  
PR 19-OCT-2000; 2000US-00692081.  
XX  
PA (MAIN-) MAINE MEDICAL CENT RES INST.  
XX  
PI Linder V, Friesel R;  
XX  
DR WPI; 2002-590472/63.  
XX N-PSDB; ABB86481.  
XX  
PT New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
XX negative remodeling, fibrosis, collagen disease, and bone disease in a  
XX mammal.

PS Claim 5; Fig 4B; 187pp; English.

XX  
XX The invention relates to mammalian REMODELIN polypeptides and encoding  
XX polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
XX negative remodeling or fibrosis, bone disease such as osteogenesis  
XX imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
XX bullosa (DEB) and Beethlem myopathy in a mammal. The polynucleotides are  
XX used to affect cellular gene expression in a mammal, where the cellular  
XX gene is from transforming growth factor (TGF)-beta1, collagen I/II/III/IV,  
XX osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein  
XX 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
XX comprising antisense REMODELIN sequences are useful for treating diseases  
XX mediated by abnormal expression of a REMODELIN molecule in a human such  
XX as impaired wound healing, fibrosis of an organ, ectopic ossification, or  
XX hypertrophic scar formation. REMODELIN is useful in the development of  
XX bone during mammalian embryogenesis, proliferation and/or migration,  
XX cellular signaling, adventitial fibrosis, negative remodeling and arterial  
XX restenosis, smooth muscle cell proliferation, and arterial remodeling and  
XX useful in gene therapy. REMODELIN is useful for elucidating the function  
XX of REMODELIN molecules in a cell, to identify a compound that affects  
XX REMODELIN expression and/or TGF-beta signaling, as a potential

CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
 CC formation, to promote bone fracture healing, and to increase or decrease  
 CC expression of REMODELIN in mammal. It is also useful for producing  
 CC recombinant cell and transgenic non-human mammals which are useful tools  
 CC for the study of REMODELIN action, for identifying novel diagnostics and  
 CC therapeutics for treatment, and for elucidating the cellular roles of  
 CC REMODELIN. The present sequence represents a human REMODELIN polypeptide  
 XX  
 SQ Sequence 243 AA;

Query Match 93.7%; Score 1225; DB 5; Length 243;  
 Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
 Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPQQLLGLFLVLLLLQLSPSSASENPKVKOKALIREVVDLYNGMCLQG 60  
 DB 1 MRPOGPAPSPQRLRG--LTLILLLLQLPAPSSASEIPKQKQAKLQREVVLDLYNGMCLQG 58  
 QY 61 PAVPGRGDSSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSSLYNGI 120  
 DB 59 PAVPGRGDSSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSSLYNGI 118  
 QY 121 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWFTFNGAECGSLPIEAIIVL 180  
 DB 119 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWFTFNGAECGSLPIEAIIVL 178  
 QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNNVSRIII 240  
 DB 179 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNNVSRIII 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243

RESULT 10  
 ABR58546  
 ID ABR58546 standard; protein; 243 AA.

AC ABR58546;  
 DT 09-JUL-2003 (first entry)

DE Human cancer related protein SEQ ID NO:203.

KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;  
 KW heart disease; atherosclerosis; endometriosiis.

OS Homo sapiens.

PN WO2003025138-A2.

PD 27-MAR-2003.

PF 17-SEP-2002; 2002WO-US029560.

PR 17-SEP-2001; 2001US-0323469P.  
 PR 20-SEP-2001; 2001US-0323687P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 08-FEB-2002; 2002US-0355145P.  
 PR 08-FEB-2002; 2002US-0355257P.  
 PR 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;  
 PI Zlotnick A;

XX MPI: 2003-354600/33.

DR N-PSDB; ACCT2666.

XX New genes that are up-regulated or down-regulated in cancers, useful as

PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as  
 XX therapeutic targets for screening drugs for treating these diseases.  
 PS Claim 12; Page 736; 767pp; English.

XX The present invention describes an isolated nucleic acid molecule, which  
 CC comprises the sequence of any of the genes that are up-regulated or down-  
 CC regulated in specific cancers (e.g. about 1011 genes up-regulated in  
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer  
 CC related gene nucleotide sequences which encode the proteins given in  
 CC ABR58521 to ABR58709. Also described: (1) determining the presence or  
 CC absence of a pathological cell in a patient; (2) an expression vector  
 CC comprising a nucleic acid molecule described above; (3) a host cell  
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by  
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide  
 CC of (4); (6) specifically targeting a compound to a pathological cell in a  
 CC patient by administering to the patient the antibody above; and (7) a  
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or  
 CC therapeutic targets. In particular, the nucleic acid is useful for  
 CC diagnosing a pathology, e.g. cancer of the bone marrow,  
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,  
 CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,  
 CC atherosclerosis and endometriosiis. The nucleic acid is also useful in  
 CC drug screening, particularly for identifying agents for treating these  
 CC pathologies

SQ Sequence 243 AA;

Query Match 93.7%; Score 1225; DB 6; Length 243;  
 Best Local Similarity 94.7%; Pred. No. 6.9e-115;  
 Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPQQLLGLFLVLLLLQLSPSSASENPKVKOKALIREVVDLYNGMCLQG 60  
 DB 1 MRPOGPAPSPQRLRG--LTLILLLLQLPAPSSASEIPKQKQAKLQREVVLDLYNGMCLQG 58  
 QY 61 PAVPGRGDSSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSSLYNGI 120  
 DB 59 PAVPGRGDSSPGANGIPGTGIPGRDGFKEGECLEESFEESWTPNYKQCSWSSLYNGI 118  
 QY 121 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWFTFNGAECGSLPIEAIIVL 180  
 DB 119 DLGKIACEFTTKMRNSALRVLFSGSLRLKCRNACCORWFTFNGAECGSLPIEAIIVL 178  
 QY 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNNVSRIII 240  
 DB 179 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGNNVSRIII 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243

RESULT 11  
 ABR48227  
 ID ABR48227 standard; protein; 243 AA.

AC ABR48227;

DT 12-JUN-2003 (first entry)

DE Human bladder cancer associated protein sequence SEQ ID NO:175.

KW Human; bladder cancer; cytostatic; gene therapy; vaccine.

XX Homo sapiens.

PN WO2003003906-A2.

PD 16-JAN-2003.

PF 03-JUL-2002; 2002WO-US021338.



QY 121 DLGKIAECTFTKRSNSALRVLPSGSLRLKCRNACCCORWYFTFNGACSGPLPIEAIITL 180  
 DB 119 DLGKIAECTFTKRSNSALRVLPSGSLRLKCRNACCCORWYFTFNGACSGPLPIEAIITL 178  
 QY 181 DQGSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRIT 240  
 DB 179 DQGSPEMNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRIT 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243  
 RESULT 13  
 ID ABUS6607 standard; protein; 243 AA.  
 AC ABUS6607;  
 XX  
 DT 02-APR-2003 (first entry)  
 XX  
 DE Lung cancer-associated polypeptide #200.  
 XX  
 KM Lung cancer-associated polypeptide; cytosolic; emphysema;  
 KM anti-inflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;  
 KM small cell lung cancer; benign lesion; precancerous lesion; bronchitis;  
 KM chronic obstructive pulmonary disease; hypersensitivity pneumonitis;  
 KM interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.  
 XX  
 OS Unidentified.  
 OS  
 PN WO200286443-A2.  
 XX  
 PD 31-OCT-2002.  
 XX  
 PF 18-APR-2002; 2002WO-US012476.  
 XX  
 PR 18-APR-2001; 2001US-0284770P.  
 PR 10-MAY-2001; 2001US-0290492P.  
 PR 09-NOV-2001; 2001US-0339245P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 29-NOV-2001; 2001US-0334370P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Aziz N, Murray R;  
 XX  
 DR WPI; 2003-093161/08.  
 DR N-PSDB; ABX76336.  
 XX  
 PT Detecting a lung cancer-associated transcript in a cell from a patient  
 PT for treating lung cancer, by contacting a biological sample from the  
 PT patient with a polynucleotide that exhibits increased or decreased  
 PT expression in lung cancer.  
 XX  
 Claim 27, Page 337, 453pp; English.  
 XX  
 CC The invention relates to a method for detecting a lung cancer-associated  
 CC transcript in a cell from a patient, comprising contacting a biological  
 CC sample from the patient with a polynucleotide that selectively hybridizes  
 CC to a sequence that is at least 80 % identical to a gene that exhibits  
 CC increased or decreased expression in lung cancer samples. Lung cancer-  
 CC associated polynucleotides and polypeptides are used for identifying a  
 CC compound that modulates a lung cancer-associated polypeptide, for  
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung  
 CC cancer in a patient and for treating a mammal having lung cancer by  
 CC administering a modulatory compound identified. The methods are useful  
 CC for treating lung cancer, such as small cell lung cancer, non-small cell  
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,  
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,  
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and  
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful

CC for diagnostic purposes and as targets for screening for therapeutic  
 CC compounds that modulate lung cancer, such as antibodies. Sequences  
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the  
 CC invention  
 XX  
 SO Sequence 243 AA;  
 Query Match 93.7%; Score 1225; DB 6; Length 243;  
 Best Local Similarity 94.7%; Pred. No. 6, 9e-115;  
 Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;  
 QY 1 MRPQGRASPOLLGLFLVLLLLDLSAPSSASENPVKVOKALIROREVVLDYNGWCLQG 60  
 DB 1 MRPQGRASPOLLGLFLVLLLLDLSAPSSASENPVKVOKALIROREVVLDYNGWCLQG 58  
 QY 61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKECECLRESFESWTPNYKQCSWSLNYGI 120  
 DB 59 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKECECLRESFESWTPNYKQCSWSLNYGI 118  
 QY 121 DLGKIAECTFTKRSNSALRVLPSGSLRLKCRNACCCORWYFTFNGACSGPLPIEAIITL 180  
 DB 119 DLGKIAECTFTKRSNSALRVLPSGSLRLKCRNACCCORWYFTFNGACSGPLPIEAIITL 178  
 QY 181 DQGSPELNTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRIT 240  
 DB 179 DQGSPEMNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRIT 238  
 QY 241 EELPK 245  
 DB 239 EELPK 243  
 RESULT 14  
 ID ADB80510 standard; protein; 243 AA.  
 AC ADB80510;  
 XX  
 DT 04-DEC-2003 (first entry)  
 XX  
 DE Ovarian cancer-associated protein #37.  
 XX  
 KM Cytosolic; gene therapy; vaccine; ovarian cancer; diagnosis;  
 KM post-operative chemotherapy; radiation therapy; tumour prognosis;  
 KM pre-cancerous lesion detection.  
 XX  
 OS Homo sapiens.  
 OS  
 PN WO2002102235-A2.  
 XX  
 PD 27-DEC-2002.  
 XX  
 PF 18-JUN-2002; 2002WO-US019297.  
 XX  
 PR 18-JUN-2001; 2001US-0299234P.  
 PR 27-AUG-2001; 2001US-0315287P.  
 PR 05-SEP-2001; 2001US-0317544P.  
 PR 13-NOV-2001; 2001US-0350666P.  
 PR 12-APR-2002; 2002US-0372246P.  
 XX  
 PA (EOSB-) EOS BIOTECHNOLOGY INC.  
 XX  
 PI Mack DH, Gish KC;  
 XX  
 DR WPI; 2003-167431/16.  
 DR N-PSDB; ADB80509.  
 XX  
 PT Detecting an ovarian cancer-associated transcript in a cell from a  
 PT patient, comprising contacting a biological sample from the patient with a  
 PT polynucleotide that hybridizes to an ovarian cancer gene.  
 XX  
 Claim 13, Page 299, 332pp; English.

The invention relates to a method of detecting an ovarian cancer-associated transcript in a cell from a patient, by contacting a biological sample from the patient with a polynucleotide that selectively hybridizes to a sequence at least 80% identical to any of one of 80 nucleic acid sequences given in the specification. The method is useful in diagnosing ovarian cancer and in identifying and using agents and/or targets that inhibit ovarian cancer. The nucleic acid molecule, polypeptide and the antibody may also be used in detecting ovarian cancer, monitoring and early detection of relapse following treatment, monitoring response to therapy, selecting patients for post-operative chemotherapy or radiation therapy, in selecting mode of therapy, determining tumour prognosis, early detection of pre-cancerous lesions, and as vaccines. This sequence corresponds to one of the proteins used for the detection method of the invention.

**SQ** Sequence 243 AA;

Query Match	93.7%	Score 1225	DB 7	Length 243
Best Local Similarity	94.7%	Pred. No. 6.9e-115		
Matches 232; Conservative	3	Mismatches 8	Indels 2	Gaps 1

[illegible]

RESULT 15  
ADN39855  
ID ADN39855 standard; protein; 243 AA.

AC ADN39855;

DT 17-JUN-2004 (first entry)

Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C225.

KM Human, differential expression; cancer; angiogenic disorder;  
KM fibrotic disorder; psoriasis; ischemia; heart disease; atherosclerosis  
KM inflammatory disease; autoimmune disease;  
KM retinal neovascularization syndrome; scarring; uterine fibroid;  
KM detection; diagnosis; prognosis; drug screening; drug targeting;  
KM wound healing; contraception; cytostatic; cardiac; immunomodulatory;  
KM vulvarey; gene therapy; vaccine

**Homo sapiens.**

PN WO2003042661-A2.

PD 22-MAY-2003.

PF 13-NOV-2002; 2002WO-US036810.

PR 13-NOV-2001; 2001US-0350666P.

PR 29-NOV-2001; 2001US-0334393P.

PR 14-DEC-2001; 2001US-0340376P.

PR	08-JAN-2002;	2002US-0347211P.
PR	10-JAN-2002;	2002US-0347345P.
PR	08-FEB-2002;	2002US-0355250P.
PR	13-FEB-2002;	2002US-0356714P.
PR	20-FEB-2002;	2002US-0359077P.
PR	29-MAR-2002;	2002US-0368809P.
PR	04-APR-2002;	2002US-0370110P.
PR	12-APR-2002;	2002US-0372246P.
PR	05-JUN-2002;	2002US-0386614P.
PR	16-JUL-2002;	2002US-0396839P.
PR	22-JUL-2002;	2002US-0397775P.
PR	22-JUL-2002;	2002US-0397845P.
PR	09-SEP-2002;	2002US-0409450P.
XX		
PA	(EOSB - ) EOS BIOTECHNOLOGY INC.	

PA (EOSB-) EOS BIOTECHNOLOGY INC.

PI Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;

XX

DR N-PSDB; ADN39637.

### PT Determining the R

PT a nucleic acid in a biological sample.

PS Claim 12; SEQ ID NO C225; 1385bp; English

CC The invention relates to nucleic acids and proteins (ADN38683-ADN40064)  
CC whose expression is upregulated or downregulated in specific cancers or  
CC other diseases such as angiogenic or fibrotic disorders, and to methods  
CC of determining the presence or absence of a pathological cell in a  
CC patient by detecting a nucleic acid at least 80% identical to those of  
CC the invention or by detecting a polypeptide of the invention. The  
CC invention also relates to expression vectors and host cells comprising a  
CC nucleic acid of the invention; antibodies which specifically bind a  
CC polypeptide of the invention; use of such antibodies for drug targeting;  
CC and methods of screening for modulators of activity or expression of the  
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,  
CC antibodies and methods are useful for diagnosing, prognosing and treating  
CC cancer and other conditions such as psoriasis, leukaemia, heart disease,  
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal  
CC neovascularisation syndromes, scarring and uterine fibroids. They may  
CC also be useful in wound healing and in contraception. The present  
CC sequence represents a polypeptide of the invention.

Sequence 243 AA:

Query Match	93.7%;	Score 1225;	DB 7;	Length 243;
Best Local Similarity	94.7%;	Pred. No. 6.9e-115;		
Matches 232; Conservative	3;	Mismatches 8;	Indels 2;	Gaps 1

QY	1	MHPGRRASPQILLGLFLVLLILLILLOLSAPSSASSENPXYKKQKLLIQREYVDVLYNMGCLQG	60
Db	1	MRPQPPASPQLRRG-LLILLILLQLPAPSSASELPKGGQKQKQLRQREYVDVLYNMGCLQG	58
QY	61	PAGVGRDGSFGANGIPGTGPIGPRDGRKGEGKCEGLRASFEEBSWTPNTYKQCSWSSLANGI	120
Db	59	PAGVGRDGSFGANGIPGTGPIGPRDGRKGEGKCEGLRASFEEBSWTPNTYKQCSWSSLANGI	118
QY	121	DLGKIACCTPFTRKMSNLSARVLFSSGLLAKCNACCQRMYPFNGAGECGSLPIEIAIITL	180
Db	119	DLGKIACCTPFTRKMSNLSARVLFSSGLSLKCRNACCQRMYPFNGAECGSLPIEIAIITL	178
QY	181	DQSPDLNSTINIHRTSSVEGCEGIGAGLVDAIVAIWGTGCSDPYKGDASTGNNVSRIIT	240
Db	179	DQSPDLMNSTINIHRTSSVEGCEGIGAGLVDAIVAIWGTGCSDPYKGDASTGNNVSRIIT	238
QY	241	EELPK 245	
Db	239	EELPK 243	

Thu Aug 4 05:41:41 2005

us-10-634-108-2.rag

Page 11

Search completed: August 4, 2005, 04:16:36  
Job time : 98.4771 secs

---

This Page Blank (uspia)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 4, 2005, 04:21:44 ; Search time 86.7909 Seconds  
(without alignments)  
1099.816 Million cell updates/sec

Title: US-10-634-108-2

Perfect score: 1307  
Sequence: 1 MHPOGRASFPOLLGLGLFLVH.....GDASTGMSVSRILIEELK 245

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

Published Applications AA:\*

1: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
3: /cgn2\_6/prodata/2/pubppa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/prodata/2/pubppa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/prodata/2/pubppa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/prodata/2/pubppa/US07\_PUBCOMB.pep:\*  
7: /cgn2\_6/prodata/2/pubppa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/prodata/2/pubppa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/prodata/2/pubppa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/prodata/2/pubppa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/prodata/2/pubppa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
16: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
17: /cgn2\_6/prodata/2/pubppa/US10\_PUBCOMB.pep:\*  
18: /cgn2\_6/prodata/2/pubppa/US10\_NEW\_PUB.pep:\*  
19: /cgn2\_6/prodata/2/pubppa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/prodata/2/pubppa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/prodata/2/pubppa/US11\_PUBCOMB.pep:\*  
22: /cgn2\_6/prodata/2/pubppa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1307	100.0	245	US-10-045-992-2	Sequence 2, Appl1
2	1307	100.0	245	US-10-634-108-2	Sequence 2, Appl1
3	1307	100.0	245	US-10-939-233-2	Sequence 2, Appl1
4	1307	100.0	277	US-10-045-992-5	Sequence 5, Appl1
5	1307	100.0	277	US-10-634-108-5	Sequence 5, Appl1
6	1307	100.0	277	US-10-939-233-5	Sequence 5, Appl1
7	1307	93.7	243	US-09-829-472A-19	Sequence 19, Appl1
8	1307	93.7	243	US-10-045-992-4	Sequence 4, Appl1
9	1307	93.7	243	US-10-097-340-135	Sequence 135, App
10	1307	93.7	243	US-10-295-027-50	Sequence 50, Appl
11	1307	93.7	243	US-10-295-027-1173	Sequence 1173, Ap

12	1225	93.7	243	15	US-10-173-999-74	Sequence 74, Appl
13	1225	93.7	243	15	US-10-058-270A-38	Sequence 38, Appl
14	1225	93.7	243	15	US-10-634-108-4	Sequence 4, Appl1
15	1225	93.7	243	15	US-10-188-832-175	Sequence 175, App
16	1225	93.7	243	18	US-10-939-233-4	Sequence 4, Appl1
17	1225	93.7	278	14	US-10-097-340-131	Sequence 131, Appl
18	1225	93.7	278	14	US-10-177-293-496	Sequence 496, App
19	1225	93.7	278	14	US-10-301-822-77	Sequence 77, Appl
20	1225	93.7	278	15	US-10-296-115-1261	Sequence 1261, Ap
21	1225	93.7	278	18	US-10-961-139-2	Sequence 2, Appl1
22	1222	93.5	243	15	US-10-351-334-205	Sequence 205, App
23	1222	93.5	278	14	US-10-060-036-4551	Sequence 4551, App
24	1220	93.3	243	14	US-10-176-847-14	Sequence 14, Appl
25	1216	93.0	243	9	US-09-834-759-514	Sequence 514, App
26	1216	93.0	243	9	US-09-938-418-7	Sequence 352, App
27	1216	93.0	243	10	US-09-946-374-352	Sequence 431, App
28	1216	93.0	243	10	US-09-946-374-431	Sequence 122, App
29	1216	93.0	243	13	US-10-006-867-122	Sequence 366, App
30	1216	93.0	243	13	US-10-052-586-366	Sequence 514, App
31	1216	93.0	243	13	US-10-007-805-514	Sequence 122, App
32	1216	93.0	243	13	US-10-063-547-122	Sequence 122, App
33	1216	93.0	243	13	US-10-063-551-122	Sequence 366, App
34	1216	93.0	243	14	US-10-174-590-366	Sequence 366, App
35	1216	93.0	243	14	US-10-176-758-366	Sequence 366, App
36	1216	93.0	243	14	US-10-175-737-366	Sequence 366, App
37	1216	93.0	243	14	US-10-063-616-122	Sequence 122, App
38	1216	93.0	243	14	US-10-174-581-366	Sequence 366, App
39	1216	93.0	243	14	US-10-176-483-366	Sequence 366, App
40	1216	93.0	243	14	US-10-176-749-366	Sequence 366, App
41	1216	93.0	243	14	US-10-176-914-366	Sequence 366, App
42	1216	93.0	243	14	US-10-176-915-366	Sequence 122, App
43	1216	93.0	243	14	US-10-063-551-122	Sequence 122, App
44	1216	93.0	243	14	US-10-063-513-122	Sequence 122, App
45	1216	93.0	243	14	US-10-063-515-122	Sequence 122, App

## ALIGNMENTS

RESULT 1		US-10-045-992-2		Application US/10045992	
		Sequence 2, Appl			
		Publication No. US2002016121A1			
		GENERAL INFORMATION:			
		APPLICANT: LINDNER, Volhard			
		APPLICANT: FRIEDEL, Robert E.			
		TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN			
		FILE REFERENCE: 053689-5006-01			
		CURRENT APPLICATION NUMBER: US/10/045,992			
		CURRENT FILING DATE: 2001-10-19			
		PRIOR APPLICATION NUMBER: US 09/692,081			
		PRIOR FILING DATE: 2000-10-19			
		NUMBER OF SEQ ID NOS: 9			
		SOFTWARE: PatentIn version 3.1			
		SEQ ID NO 2			
		LENGTH: 245			
		TYPE: PRT			
		ORGANISM: Rattus sp.			
		US-10-045-992-2			
Query Match		100.0%		Score 1307; DB 13; Length 245;	
Best Local Similarity		100.0%		Pred. No. 4.7e-121;	
Matches 245; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
QY		1 MHPOGRASFPOLLGLGLFLVLLLOLSASASSASBNPKVKOKALIRQREVVLDYNGKCLQG 60			
Db		1 MHPOGRASFPOLLGLGLFLVLLLOLSASASSASBNPKVKOKALIRQREVVLDYNGKCLQG 60			
QY		61 PAVGVRDSSPANGIGPTGPIRGPDGFGKGECELRBSFESSWPNYQCSWSSLYNGI 120			
Db		61 PAVGVRDSSPANGIGPTGPIRGPDGFGKGECELRBSFESSWPNYQCSWSSLYNGI 120			
QY		121 DGGKIAECTFTMRNSALRVLFSGSLRLKCNACCQRRVYFPFNGACGSPLEIAIIVL 180			

```
Db      121 |||||
        DGGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCCORWYFTFNGAECGSPLEIAIYYL 180
QY      181 DGGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIII 240
        |||||
        DGGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIII 240
Db      241 EELPK 245
        |||||
        241 EELPK 245

RESULT 2
US-10-634-108-2
; Sequence 2, Application US/10634108
; Publication No. US2004063176A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/10/634,108
; PRIOR FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-634-108-2

Query Match      100.0%; Score 1307; DB 15; Length 245;
Best Local Similarity 100.0%; Pred. No. 4.7e-121;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MHPOGRAASPOLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGKCLQG 60
        1 MHPOGRAASPOLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGKCLQG 60
Db      61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKEGECRLRESFESWTPNYKQCSWSSLNYGI 120
        61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKEGECRLRESFESWTPNYKQCSWSSLNYGI 120
Db      121 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCCORWYFTFNGAECGSPLEIAIYYL 180
        |||||
        DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCCORWYFTFNGAECGSPLEIAIYYL 180
QY      181 DGGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIII 240
        |||||
        DGGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIII 240
Db      241 EELPK 245
        |||||
        241 EELPK 245

RESULT 3
US-10-939-233-2
; Sequence 2, Application US/10939233
; Publication No. US20050147602A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTRHCL, A NOVEL,
; FILE REFERENCE: 053689-5006-03
; CURRENT APPLICATION NUMBER: US/10/939,233
; CURRENT FILING DATE: 2004-09-10
; PRIOR APPLICATION NUMBER: US 09/692,081
; PRIOR FILING DATE: 2000-10-19
; PRIOR APPLICATION NUMBER: US 10/045,992
```

```
; PRIOR FILING DATE: 2001-10-19
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-939-233-2
```

```
Query Match      100.0%; Score 1307; DB 18; Length 245;
Best Local Similarity 100.0%; Pred. No. 4.7e-121;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MHPOGRAASPOLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGKCLQG 60
        1 MHPOGRAASPOLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGKCLQG 60
Db      61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKEGECRLRESFESWTPNYKQCSWSSLNYGI 120
        61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKEGECRLRESFESWTPNYKQCSWSSLNYGI 120
Db      121 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCCORWYFTFNGAECGSPLEIAIYYL 180
        |||||
        DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCCORWYFTFNGAECGSPLEIAIYYL 180
QY      181 DGGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIII 240
        |||||
        DGGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIII 240
Db      241 EELPK 245
        |||||
        241 EELPK 245
```

```
RESULT 4
US-10-045-992-5
; Sequence 5, Application US/10045992
; Publication No. US20020161211A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert E.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN
; FILE REFERENCE: 053689-5006-01
; CURRENT APPLICATION NUMBER: US/10/045,992
; CURRENT FILING DATE: 2001-10-19
; PRIOR APPLICATION NUMBER: US 09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-045-992-5
```

```
Query Match      100.0%; Score 1307; DB 13; Length 277;
Best Local Similarity 100.0%; Pred. No. 5.5e-121;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MHPOGRAASPOLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGKCLQG 60
        33 MHPOGRAASPOLLGFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGKCLQG 92
Db      61 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKEGECRLRESFESWTPNYKQCSWSSLNYGI 120
        93 PAGVPGRDSPGANGIPGTPGIPGRDGFKEGKEGECRLRESFESWTPNYKQCSWSSLNYGI 152
Db      121 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCCORWYFTFNGAECGSPLEIAIYYL 180
        153 DLGKIAECTFTKRRNSALRVLFSSGLRLKCRNACCCORWYFTFNGAECGSPLEIAIYYL 212
QY      181 DGGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPKGDASTGNNVSRIII 240
        |||||
```

Db 213 DQSPPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSR111 272  
QY 241 EELPK 245  
273 EELPK 277

RESULT 5  
US-10-634-108-5  
; Sequence 5, Application US/10634108  
; Publication No. US20040063176A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volhard  
; APPLICANT: PRIESEL, Robert F.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL  
; FILE REFERENCE: 36-13 Lindner et al. (20036.0013)  
; CURRENT APPLICATION NUMBER: US/10/634,108  
; PRIOR FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/692,081  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-634-108-5

Query Match 100.0%; Score 1307; DB 15; Length 277;  
Best Local Similarity 100.0%; Pred. No. 5, 5e-121;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPQILLGLFVLLILLQLSAPSSASENPKVKOKALIRORREVVLDYNGCLOG 60  
33 MHPOGRAASPQILLGLFVLLILLQLSAPSSASENPKVKOKALIRORREVVLDYNGCLOG 92  
Db 61 PAVPGRDSSPGANGIPGTPGIPGRDGFGEKGECLRESFEESMTPNYKQCSWSSINYG1 120  
93 PAVPGRDSSPGANGIPGTPGIPGRDGFGEKGECLRESFEESMTPNYKQCSWSSINYG1 152  
QY 121 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNAGCSGPLEIAI1YL 180  
153 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNAGCSGPLEIAI1YL 212  
Db 181 DQSPPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSR111 240  
213 DQSPPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSR111 272  
QY 241 EELPK 245  
273 EELPK 277  
Db

RESULT 6  
US-10-939-233-5  
; Sequence 5, Application US/10939233  
; Publication No. US20050147602A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volhard  
; APPLICANT: PRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTRHC1, A NOVEL  
; FILE REFERENCE: 053689-5006-03  
; CURRENT APPLICATION NUMBER: US/10/939,233  
; PRIOR FILING DATE: 2004-09-10  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US 10/045,992  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 277

; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-939-233-5

Query Match 100.0%; Score 1307; DB 18; Length 277;  
Best Local Similarity 100.0%; Pred. No. 5, 5e-121;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MHPOGRAASPQILLGLFVLLILLQLSAPSSASENPKVKOKALIRORREVVLDYNGCLOG 60  
33 MHPOGRAASPQILLGLFVLLILLQLSAPSSASENPKVKOKALIRORREVVLDYNGCLOG 92  
Db 61 PAVPGRDSSPGANGIPGTPGIPGRDGFGEKGECLRESFEESMTPNYKQCSWSSINYG1 120  
93 PAVPGRDSSPGANGIPGTPGIPGRDGFGEKGECLRESFEESMTPNYKQCSWSSINYG1 152  
QY 121 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNAGCSGPLEIAI1YL 180  
153 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNAGCSGPLEIAI1YL 212  
Db 181 DQSPPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSR111 240  
213 DQSPPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSR111 272  
QY 241 EELPK 245  
273 EELPK 277  
Db

RESULT 7  
US-09-829-472A-19  
; Sequence 19, Application US/09829472A  
; Publication No. US20040146862A1  
; GENERAL INFORMATION:  
; APPLICANT: Mack, David  
; APPLICANT: Gish, Kurt  
; APPLICANT: EOS Biotechnology, Inc.  
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer, Compositions and  
; FILE REFERENCE: 018501-001200US  
; CURRENT APPLICATION NUMBER: US/09/829,472A  
; PRIOR FILING DATE: 2001-04-09  
; PRIOR APPLICATION NUMBER: US 09/525,361  
; NUMBER OF SEQ ID NOS: 20  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 19  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: BCN4, ESTs, secreted protein  
US-09-829-472A-19

Query Match 93.7%; Score 1225; DB 12; Length 243;  
Best Local Similarity 94.7%; Pred. No. 6, 3e-113;  
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPQILLGLFVLLILLQLSAPSSASENPKVKOKALIRORREVVLDYNGCLOG 60  
1 MHPOGRAASPQILLGLFVLLILLQLSAPSSASENPKVKOKALIRORREVVLDYNGCLOG 58  
Db 61 PAVPGRDSSPGANGIPGTPGIPGRDGFGEKGECLRESFEESMTPNYKQCSWSSINYG1 120  
59 PAVPGRDSSPGANGIPGTPGIPGRDGFGEKGECLRESFEESMTPNYKQCSWSSINYG1 118  
QY 121 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNAGCSGPLEIAI1YL 180  
119 DLGKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORRYFTFNAGCSGPLEIAI1YL 178  
Db 181 DQSPPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSR111 240  
179 DQSPPELNTSTINIHRTSSVGLCEGIGAGLVDAIIVGTCSDPYKGDASTGMSVSR111 238



```
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 50
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-50
```

```
Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 222; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHQGRASAPQLLGLFLVLLLLQLSAPSSASENPVKKALIRQREVVDLYNGCLOG 60
DB 1 MRQGPASAPQRLRG--LTLTLLLLQLPAPSSASEIPKQKQALRQREVVDLYNGCLOG 58
QY 61 PAVPGRDGSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSWSLNYGI 120
DB 59 PAVPGRDGSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSWSLNYGI 118
QY 121 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 180
DB 119 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 178
QY 181 DQSPPELNTSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 240
DB 179 DQSPPELNTSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 238
QY 241 EELPK 245
DB 239 EELPK 243
```

## RESULT 11

```
US-10-295-027-1173
; Sequence 1173, Application US/10295027
; Publication NO. US2003022350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natascha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glyme, Richard
; APPLICANT: Hevez, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-012500US
; PRIOR APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
```

```
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1173
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-1173
```

```
Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 222; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHQGRASAPQLLGLFLVLLLLQLSAPSSASENPVKKALIRQREVVDLYNGCLOG 60
DB 1 MRQGPASAPQRLRG--LTLTLLLLQLPAPSSASEIPKQKQALRQREVVDLYNGCLOG 58
QY 61 PAVPGRDGSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSWSLNYGI 120
DB 59 PAVPGRDGSPGANGIPGTGPIGRDGFGEKGECECLRESPEESWTPNYKQCSWSLNYGI 118
QY 121 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 180
DB 119 DGLKIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTFNGAECGSLPIEAIIVL 178
QY 181 DQSPPELNTSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 240
DB 179 DQSPPELNTSTINIHRTSVYGLCEGIGAGLVDAIIVWGTCSDPKGDASTGNNVSRIII 238
QY 241 EELPK 245
DB 239 EELPK 243
```

## RESULT 12

```
US-10-173-999-74
; Sequence 74, Application US/10173999
; Publication NO. US2004000563A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Ovarian Cancer, Compositions
; FILE REFERENCE: 018501-002420US
; PRIOR APPLICATION NUMBER: US/10/173,999
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: US 60/299,234
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/315,287
; PRIOR FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: US 60/350,666
```

```
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2001-04-12
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-173-999-74

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPQULLLGLFVLLLLQLSAPSSASENPVKOKALIRQREVVDLYNMCLOG 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRPOGPASPORLRG--LTLTLLTLLQLPAPSSASEIPKQKQKQLRQREVVDLYNMCLOG 58
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PAVVPGRDGSPGANGIPGTGGIPGRDGFKEGKEGECLEBSFESWTPNYKQCSWSSLYNGI 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 59 PAVVPGRDGSPGANGIPGTGGIPGRDGFKEGKEGECLEBSFESWTPNYKQCSWSSLYNGI 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNAGCSGPLEIAIYYL 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 119 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNAGCSGPLEIAIYYL 178
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 DQGSPELNTSTNIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSR111 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 179 DQGSPEMNSTINIHRITSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSR111 238
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 EELPK 245
   |||||
Db 239 EELPK 243
   |||||

RESULT 13
US-10-058-270A-38
; Sequence 38, Application US/10058270A
; Publication No. US20040029114A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Afar, Daniel
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Breast Cancer Compositions and
; FILE REFERENCE: 018501-005210US
; CURRENT APPLICATION NUMBER: US/10/058,270A
; CURRENT FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: US 60/263,965
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: US 60/265,928
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: US 09/829,472
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/282,698
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 60/288,590
; PRIOR FILING DATE: 2001-05-04
; PRIOR APPLICATION NUMBER: US 60/294,443
; PRIOR FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 141
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-058-270A-38

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```
QY 1 MHPOGRAASPQULLLGLFVLLLLQLSAPSSASENPVKOKALIRQREVVDLYNMCLOG 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRPOGPASPORLRG--LTLTLLTLLQLPAPSSASEIPKQKQKQLRQREVVDLYNMCLOG 58
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PAVVPGRDGSPGANGIPGTGGIPGRDGFKEGKEGECLEBSFESWTPNYKQCSWSSLYNGI 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 59 PAVVPGRDGSPGANGIPGTGGIPGRDGFKEGKEGECLEBSFESWTPNYKQCSWSSLYNGI 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNAGCSGPLEIAIYYL 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 119 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNAGCSGPLEIAIYYL 178
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 DQGSPELNTSTNIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSR111 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 179 DQGSPEMNSTINIHRITSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSR111 238
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 EELPK 245
   |||||
Db 239 EELPK 243
   |||||

RESULT 14
US-10-634-108-4
; Sequence 4, Application US/10634108
; Publication No. US20040063176A1
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/10/634,108
; CURRENT FILING DATE: 2003-08-04
; PRIOR APPLICATION NUMBER: US/09/692,081
; PRIOR FILING DATE: 2000-10-19
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-634-108-4

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 1 MHPOGRAASPQULLLGLFVLLLLQLSAPSSASENPVKOKALIRQREVVDLYNMCLOG 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MRPOGPASPORLRG--LTLTLLTLLQLPAPSSASEIPKQKQKQLRQREVVDLYNMCLOG 58
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 61 PAVVPGRDGSPGANGIPGTGGIPGRDGFKEGKEGECLEBSFESWTPNYKQCSWSSLYNGI 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 59 PAVVPGRDGSPGANGIPGTGGIPGRDGFKEGKEGECLEBSFESWTPNYKQCSWSSLYNGI 118
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNAGCSGPLEIAIYYL 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 119 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCORWYFTFNAGCSGPLEIAIYYL 178
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 181 DQGSPELNTSTNIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSR111 240
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 179 DQGSPEMNSTINIHRITSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSR111 238
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

QY 241 EELPK 245
   |||||
Db 239 EELPK 243
   |||||

RESULT 15
US-10-188-832-175
; Sequence 175, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natesha
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 175
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-175
```

```

Query Match          93.7%; Score 1225; DB 15; Length 243;
Best Local Similarity 94.7%; Pred. No. 6.3e-113;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;
```

```

QY      1 MHPQGRASRQLLGLFVYLLLLQLSAPSSASENPKYKALIRQREVVDLYNGMCLQG 60
      1 MRQGPASRQRLRG--LTLTLLQLPAPSSASEIPKQKQALRQREVVDLYNGMCLQG 58
QY      61 PAGVPERDSDPGANGIPGTPGIPGRDGFKEGKEGCELRSPFESMTPNYKQCSWSLNYGI 120
      59 PAGVPERDSDPGANGIPGTPGIPGRDGFKEGKEGCELRSPFESMTPNYKQCSWSLNYGI 118
QY      121 DLGKIAECTFTKRSNSALRVLFSSGLRLKCRNACCORWYFTFNAGCSGPLEIAIIVL 180
      119 DLGKIAECTFTKRSNSALRVLFSSGLRLKCRNACCORWYFTFNAGCSGPLEIAIIVL 178
QY      181 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPKGDASTGANSVSRIII 240
      179 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIVWGTCSDPKGDASTGANSVSRIII 238
QY      241 EELPK 245
      239 EELPK 243
Db
```

```

Search completed: August 4, 2005, 04:45:26
Job time : 89.7909 secs
```

This Page Blank (uspio)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:56 ; Search time 27.881 Seconds

(Without alignments)  
741.643 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAELGQTLSPAGLCRPL.....GDASTGMSVSRITIEELPK 277

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patente\_AA:\*  
1: /cgn2\_6/prodata/1/aa/5A.COMB.pep:\*  
2: /cgn2\_6/prodata/1/aa/5B.COMB.pep:\*  
3: /cgn2\_6/prodata/1/aa/6A.COMB.pep:\*  
4: /cgn2\_6/prodata/1/aa/6B.COMB.pep:\*  
5: /cgn2\_6/prodata/1/aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/prodata/1/aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1476	100.0	277	4	US-09-692-081-5
2	1307	88.6	245	4	US-09-692-081-2
3	1234.5	83.6	278	4	US-09-834-759-515
4	1225	83.0	243	4	US-09-692-081-4
5	1222	82.8	243	4	US-09-489-847-205
6	1216	82.4	243	4	US-09-834-759-514
7	1103.5	74.8	232	4	US-09-205-258-958
8	986.5	66.8	232	4	US-09-834-759-517
9	968	65.6	197	4	US-09-834-759-516
10	353	23.9	66	4	US-09-205-258-962
11	296	20.1	52	4	US-09-205-258-961
12	270	18.3	51	4	US-09-205-258-963
13	248	16.8	46	4	US-09-834-759-518
14	215	14.6	52	4	US-09-205-258-960
15	141	9.6	26	4	US-09-834-759-519
16	130	8.8	93	4	US-09-489-847-240
17	130	8.8	93	4	US-09-489-847-362
18	125	8.5	63	3	US-09-029-348-3
19	125	8.5	66	3	US-09-029-348-2
20	124.5	8.4	1516	3	US-09-949-016-8209
21	121.5	8.2	128	3	US-09-227-357-190
22	121	8.2	755	4	US-09-919-497-57
23	119.5	8.1	246	2	US-08-463-911-4
24	119.5	8.1	313	4	US-09-949-016-9265
25	119.5	8.1	1670	4	US-09-949-016-5883
26	119	8.1	245	4	US-09-552-225A-4
27	119	8.1	245	4	US-09-552-204A-4

28	117	7.9	96	4	US-09-513-999C-4206	Sequence 4206, Ap
29	117	7.9	96	4	US-09-471-276-823	Sequence 823, Ap
30	117	7.9	208	4	US-09-800-729-151	Sequence 151, Ap
31	117	7.9	1057	3	US-08-931-820-4	Sequence 4, Appl
32	117	7.9	1078	3	US-08-963-825-21	Sequence 21, Appl
33	117	7.9	1078	3	US-09-500-811-21	Sequence 21, Appl
34	117	7.9	1078	3	US-09-570-573-21	Sequence 21, Appl
35	117	7.9	1078	3	US-09-548-608-21	Sequence 21, Appl
36	116	7.9	258	4	US-09-976-594-815	Sequence 815, Ap
37	116	7.9	744	4	US-09-949-016-9607	Sequence 9607, Ap
38	115.5	7.8	349	6	US-10466-2	Patent No. 5510466
39	115.5	7.8	349	6	5510466-2	Patent No. 5510466
40	115.5	7.8	453	6	5510466-4	Patent No. 5510466
41	115.5	7.8	453	6	5510466-4	Patent No. 5510466
42	115.5	7.8	684	1	US-08-555-669-12	Sequence 12, Appl
43	115.5	7.8	684	1	US-09-073-663-12	Sequence 12, Appl
44	115	7.8	1078	4	US-09-949-016-1185	Sequence 1185, A
45	115	7.8	1143	4	US-09-949-016-6137	Sequence 6137, Ap

## ALIGNMENTS

```
RESULT 1
US-09-692-081-5
; Sequence 5, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhart
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al (2000036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-692-081-5

Query Match      100.0%; Score 1476; DB 4; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.3e-150;
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPAELGQTLSPAGLCRPLCLLCSQUPHTWHPGGRASPQLLGLFLVLLLLQLSA 60
      |||
Db      1 MRPAELGQTLSPAGLCRPLCLLCSQUPHTWHPGGRASPQLLGLFLVLLLLQLSA 60
      |||

QY      61 PSSASNPVKYKQALIRORRYVDLYNGMCLQGPAGVPRGDSFGANGICITPTPTIPRDRG 120
      |||
Db      61 PSSASNPVKYKQALIRORRYVDLYNGMCLQGPAGVPRGDSFGANGICITPTPTIPRDRG 120
      |||

QY      121 KGEKGECLSEFSESTPYKQCSMSLNYGIDLGIACTPTPKMSNSALRYLFGSLR 180
      |||
Db      121 KGEKGECLSEFSESTPYKQCSMSLNYGIDLGIACTPTPKMSNSALRYLFGSLR 180
      |||

QY      181 LKCRNACCRWYFTFNGACSGPLTEAITIYLDQSPBELNSTINIHRTSVBGLCGIG 240
      |||
Db      181 LKCRNACCRWYFTFNGACSGPLTEAITIYLDQSPBELNSTINIHRTSVBGLCGIG 240
      |||

QY      241 GLYDVAIWGTGSDYKPGDASTGMSVSRITIEELPK 277
      |||
Db      241 GLYDVAIWGTGSDYKPGDASTGMSVSRITIEELPK 277
      |||

RESULT 2
US-09-692-081-2
; Sequence 2, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhart
```

```
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 245
; TYPE: PRT
; ORGANISM: Rattus sp.
US-09-692-081-2

Query Match      88.6%; Score 1307; DB 4; Length 245;
Best Local Similarity 100.0%; Pred. No. 1.8e-132;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVQKALIRORREVVLDYNGMCLQG 92
DB 1 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVQKALIRORREVVLDYNGMCLQG 60
QY 93 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECIRSFESBWTPTYKQCSWSSLNYGI 152
DB 61 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECIRSFESBWTPTYKQCSWSSLNYGI 120
QY 153 DKGIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTNGACSGPLPIEAIIVL 212
DB 121 DKGIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTNGACSGPLPIEAIIVL 180
QY 213 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPRKGDASTGMNSVSRIT 272
DB 181 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPRKGDASTGMNSVSRIT 240
QY 273 EELPK 277
DB 241 EELPK 245

RESULT 3
US-09-834-759-515
; Sequence 515, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: JIANG, Yugu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: PasteSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515

Query Match      83.6%; Score 1234.5; DB 4; Length 278;
Best Local Similarity 86.4%; Pred. No. 1.4e-124;
Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;
```

```
DB 59 LPAPSSASEIPKQKQALRQREVVLDYNGMCLQGPAGVPGRDSPGANVIRPTGPIGR 118
QY 118 DGFKEGECIRSFESBWTPTYKQCSWSSLNYGIDKGIAECTFTKRSNSALRVLSG 177
DB 119 DGFKEGECIRSFESBWTPTYKQCSWSSLNYGIDKGIAECTFTKRSNSALRVLSG 178
QY 178 SLRLKCRNACCORWYFTNGACSGPLPIEAIIVLDOSPELNSTINIHRTSSVEGLCEG 237
DB 179 SLRLKCRNACCORWYFTNGACSGPLPIEAIIVLDOSPELNSTINIHRTSSVEGLCEG 238
QY 238 IGAGLVDAIIVWGTCSDYPRKGDASTGMNSVSRITIEELPK 277
DB 239 IGAGLVDAIIVWGTCSDYPRKGDASTGMNSVSRITIEELPK 278

RESULT 4
US-09-692-081-4
; Sequence 4, Application US/09692081
; Patent No. 6630325
; GENERAL INFORMATION:
; APPLICANT: LINDNER, Volkhard
; APPLICANT: FRIESEL, Robert F.
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL
; FILE REFERENCE: 36-13 Lindner et al. (200036.0013)
; CURRENT APPLICATION NUMBER: US/09/692,081
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-692-081-4

Query Match      83.0%; Score 1225; DB 4; Length 243;
Best Local Similarity 94.7%; Pred. No. 1.2e-123;
Matches 232; Conservative 3; Mismatches 8; Indels 2; Gaps 1;

QY 33 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVQKALIRORREVVLDYNGMCLQG 92
DB 1 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVQKALIRORREVVLDYNGMCLQG 58
QY 93 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECIRSFESBWTPTYKQCSWSSLNYGI 152
DB 59 PAGVPGRDSPGANGIPGTGIPGRDGFKEGECIRSFESBWTPTYKQCSWSSLNYGI 118
QY 153 DKGIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTNGACSGPLPIEAIIVL 212
DB 119 DKGIAECTFTKRSNSALRVLFSGSLRLKCRNACCORWYFTNGACSGPLPIEAIIVL 178
QY 213 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPRKGDASTGMNSVSRIT 272
DB 179 DOGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDYPRKGDASTGMNSVSRIT 238
QY 273 EELPK 277
DB 239 EELPK 243

RESULT 5
US-09-489-847-205
; Sequence 205, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: ROSEN et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; NUMBER OF SEQ ID NOS: 2000-01-24
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1998-07-30
```

```

; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319
; EARLIER FILING DATE: 1998-08-12
; EARLIER APPLICATION NUMBER: 60/095,454
; EARLIER FILING DATE: 1998-08-06
; EARLIER APPLICATION NUMBER: 60/095,455
; EARLIER FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 205
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-489-847-205

```

```

Query Match      82.4%; Score 1222; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 2.6e-123;
Matches 231; Conservative 4; Mismatches 8; Indels 2; Gaps 1;

```

```

QY 33 MHPOGSAASPOLLLGLFVLLILLQLSAPSSASENPKVKKALIRQREVVDLYNGMCTOG 92
DB 1 MRPOGPAASPORLGG--LILLILLQLPAPSSASEIPKQKQKQLRQREVVDLYNGMCTOG 58
QY 93 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESEESWTPTYKCCSSSLNYGI 152
DB 59 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESEESWTPTYKCCSSSLNYGI 118
QY 153 DLGKIACFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEAIYYL 212
DB 119 DLGKIACFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEAIYYL 178
QY 213 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 272
DB 179 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 238
QY 273 EELPK 277
DB 239 EELPK 243

```

```

RESULT 6
US-09-834-759-514
; Sequence 514, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121, 470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 514
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-514

```

```

Query Match      82.4%; Score 1216; DB 4; Length 243;
Best Local Similarity 94.3%; Pred. No. 1.1e-122;
Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

```

```

QY 33 MHPOGSAASPOLLLGLFVLLILLQLSAPSSASENPKVKKALIRQREVVDLYNGMCTOG 92
DB 1 MRPOGPAASPORLGG--LILLILLQLPAPSSASEIPKQKQKQLRQREVVDLYNGMCTOG 58

```

```

QY 93 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESEESWTPTYKCCSSSLNYGI 152
DB 59 PAVGPRDGSFGANGIPGTGIPGRDGFKEGKECECLRESEESWTPTYKCCSSSLNYGI 118
QY 153 DLGKIACFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEAIYYL 212
DB 119 DLGKIACFTTKRNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEAIYYL 178
QY 213 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 272
DB 179 DQSSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDPKGDASTGMNSVSRIT 238
QY 273 EELPK 277
DB 239 EELPK 243

```

```

RESULT 7
US-09-205-258-958
; Sequence 958, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972

```

```

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 958
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-205-258-938

```

```

Query Match          74.8%; Score 1103.5; DB 4; Length 276;
Best Local Similarity 74.9%; Pred. No. 1.8e-110;
Matches 221; Conservative 6; Mismatches 27; Indels 41; Gaps 5;

```

```

QY 2 RPAELGQTL-----SRAGLCRPLCLLCASQDPTHMPOGRAASPOLLLGLFLVLLLI 54
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 4 RPSASXQTLTTLTFLSSVSSAS-----SSALGSGRPPCDPPRAPPFRSG----- 46
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 55 LQLLSAPSSAS-----ENPKVQKALIRREVVDLYNGMCLGSPAGVGRDGS 102
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 47 ----SASCCSCCCSCPPRRAPLRSPR-GSKRIKREVVLDLNGMCLGSPAGVGRDGS 101
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 103 PGANGIPGTPIGRDGFKEKCEKCELRSEFEESWTNPKOCSSWSLNYGIDLGKIAECTF 162
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 102 PGANGIPGTPIGRDGFKEKCEKCELRSEFEESWTNPKOCSSWSLNYGIDLGKIAECTF 161
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 163 TKRNSALRVLFFSGSLRLKCRACCCORWFTFNAGACSGPLFIEAIIYIDGSPELNST 222
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 162 TKRNSALRVLFFSGSLRLKCRNACCCORWFTFNAGACSGPLFIEAIIYIDGSPENNST 221
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 223 INHRTSSVEGLCEGIGAGLVDAIVWGTCSDDPKGDASTGMSVSRITIEELPK 277
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 222 INHRTSSVEGLCEGIGAGLVDAIVWGTCSDDPKGDASTGMSVSRITIEELPK 276
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 8

```

US-09-834-759-517
; Sequence 517, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 517
; LENGTH: 232
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-517

```

```

Query Match          66.8%; Score 986.5; DB 4; Length 232;
Best Local Similarity 83.8%; Pred. No. 5.7e-98;
Matches 196; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

```

```

QY 1 MRPAEL--GOTLSRAGLCRPLCLLCASQDPTHMPOGRAASPOLLLGLFLVLLLIQ 57
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 1 MQPAASERGADADHVPILGLRLQLRAARQPGAMRPGAPASFORLNG--LTLILLIQQ 58
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 58 LSAPSSASENPKVQKALIRREVVDLYNGMCLGSPAGVGRDGS PGANGIPGTGIPGR 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 59 LPAPSSASERPGKKAQDLRREVVDLYNGMCLGSPAGVGRDGS PGANVLPGTGIPGR 118
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 118 DGFKEKCEKCELRSEFEESWTNPKOCSSWSLNYGIDLGKIAECTFTKRNSALRVLPSG 177
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 119 DGFKEKCEKCELRSEFEESWTNPKOCSSWSLNYGIDLGKIAECTFTKRNSALRVLPSG 178
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 178 SRLKCRNACCCORWFTFNAGACSGPLFIEAIIYIDGSPELNSTINIHRTSSV 231
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 179 SRLKCRNACCCORWFTFNAGACSGPLFIEAIIYIDGSPENNSTINIHRTSSV 232
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

RESULT 9

```

US-09-834-759-516
; Sequence 516, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugu
; APPLICANT: Dillon, David C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 516
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-516

```

```

Query Match          65.6%; Score 968; DB 4; Length 197;
Best Local Similarity 93.0%; Pred. No. 4.4e-96;
Matches 185; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

```



EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,916  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,373  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,875  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,374  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,917  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,949  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,974  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,883  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,897  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,898  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,962  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,963  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,877  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,878  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/070,923  
EARLIER FILING DATE: 1997-12-18  
EARLIER APPLICATION NUMBER: 60/092,921  
EARLIER FILING DATE: 1998-07-15  
EARLIER APPLICATION NUMBER: 60/094,657  
EARLIER FILING DATE: 1998-07-30

NUMBER OF SEQ ID NOS: 1227  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 961  
LENGTH: 52  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-205-258-961  
  
Query Match 20.1% Score 296; DB 4; Length 52;  
Best Local Similarity 100.0%; Pred. No. 1,7e-24;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 111 TPGI PRDGFGEKGEKCELRSEFSFESWTPTNYKQCSWSLNYGIDGKIACTP 162  
Db 1 TPGI PRDGFGEKGEKCELRSEFSFESWTPTNYKQCSWSLNYGIDGKIACTP 52  
  
RESULT 12  
US-09-205-258-963  
Sequence 963, Application US/09205258  
Patent No. 6525174  
GENERAL INFORMATION:  
APPLICANT: Young et al.  
TITLE OF INVENTION: 207 Human Secreted Proteins  
FILE REFERENCE: P2007P1  
CURRENT APPLICATION NUMBER: US/09/205,258  
CURRENT FILING DATE: 1998-12-04  
EARLIER APPLICATION NUMBER: FCT/US98/11422  
EARLIER FILING DATE: 1998-06-04  
EARLIER APPLICATION NUMBER: 60/048,885  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,375  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,881  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,880  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,896  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,020  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,876  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,895  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,884  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,894  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,971  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,964  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,882  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,899  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,893  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,900  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,901  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,892  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,915  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/049,019  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,970  
EARLIER FILING DATE: 1997-06-06  
EARLIER APPLICATION NUMBER: 60/048,972  
EARLIER FILING DATE: 1997-06-06

```

; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 963
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-205-258-963

Query Match      18.3%; Score 270; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      227 RTSSVGLCEGIGAGLVDAIWTGTCSDYKGDASTGMNSVSRILIELPK 277
Db      1 RTSSVGLCEGIGAGLVDAIWTGTCSDYKGDASTGMNSVSRILIELPK 51

RESULT 13
; US-09-834-759-518
; Sequence 518, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 518
; LENGTH: 46
; TYPE: PRT
```

```

; ORGANISM: Homo sapiens
; US-09-834-759-518

Query Match      16.8%; Score 248; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.1e-19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      232 EGICEGIGAGLVDAIWTGTCSDYKGDASTGMNSVSRILIELPK 277
Db      1 EGICEGIGAGLVDAIWTGTCSDYKGDASTGMNSVSRILIELPK 46

RESULT 14
; US-09-205-258-960
; Sequence 960, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentn Ver. 2.0
; SEQ ID NO 960
; LENGTH: 52
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-960
```

```
Query Match 14.6%; Score 215; DB 4; Length 52;
Best Local Similarity 95.2%; Pred. No. 8.9e-16;
Matches 40; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 73 KALIRGQEVVDLYNGMCLQGPGAGVPGRDSPGANGIPGTPI 114
| | | | | | | | | | | | | | | | | | | | | | | |
Db 11 KRIRGQEVVDLYNGMCLQGPGAGVPGRDSPGANGIPGTPI 52
```

```
RESULT 15
US-09-834-759-519
; Sequence 519, Application US/09834759
; Patent No. 6680197
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yugui
; APPLICANT: Dillon, Devin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jianshun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834,759
; CURRENT FILING DATE: 2001-04-13
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: PatcSeq for Windows Version 3.0
; SEQ ID NO 519
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-519
```

```
Query Match 9.6%; Score 141; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 3e-08;
```

```
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 252 CSDYPRKGDASTGWNVSRIIIIEELPK 277
| | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CSDYPRKGDASTGWNVSRIIIIEELPK 26
```

```
Search completed: August 4, 2005, 04:23:54
Job time : 28.881 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:21:44 ; Search time 98.1268 Seconds  
(without alignments)  
1099.816 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAELGQTLSRAGLCRPL.....GDASTGMSVSNIIIELEPK 277

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1745140 seqs, 389608008 residues

Total number of hits satisfying chosen parameters: 1745140

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10E\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10F\_PUBCOMB.pep:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US11\_PUBCOMB.pep:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep:\*  
21: /cgn2\_6/ptodata/2/pubpaa/US11\_PUBCOMB.pep:\*  
22: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1476	100.0	277	US-10-045-992-5	Sequence 5, Appli
2	1476	100.0	277	US-10-634-108-5	Sequence 5, Appli
3	1476	100.0	277	US-10-939-233-5	Sequence 5, Appli
4	1307	88.6	245	US-10-045-992-2	Sequence 2, Appli
5	1307	88.6	245	US-10-634-108-2	Sequence 2, Appli
6	1307	88.6	245	US-10-939-233-2	Sequence 2, Appli
7	1243.5	84.2	278	US-10-097-340-111	Sequence 131, App
8	1243.5	84.2	278	US-10-177-293-496	Sequence 496, App
9	1243.5	84.2	278	US-10-301-822-77	Sequence 77, Appli
10	1243.5	84.2	278	US-10-961-139-2	Sequence 2, Appli
11	1240.5	84.0	278	US-10-060-036-4551	Sequence 4551, Ap

12	1239.5	84.0	278	15	US-10-296-115-1261	Sequence 1261, Ap
13	1234.5	83.6	278	9	US-09-834-759-515	Sequence 515, App
14	1234.5	83.6	278	13	US-10-007-805-515	Sequence 515, App
15	1234.5	83.6	278	14	US-10-076-622-515	Sequence 515, App
16	1234.5	83.6	278	14	US-10-060-036-4560	Sequence 4560, Ap
17	1234.5	83.6	278	14	US-10-124-805-515	Sequence 515, App
18	1225	83.0	243	12	US-09-829-472A-19	Sequence 19, Appli
19	1225	83.0	243	13	US-10-045-992-4	Sequence 4, Appli
20	1225	83.0	243	14	US-10-097-340-135	Sequence 135, App
21	1225	83.0	243	15	US-10-295-027-50	Sequence 50, Appli
22	1225	83.0	243	15	US-10-173-999-74	Sequence 173, Ap
23	1225	83.0	243	15	US-10-058-270A-38	Sequence 74, Appli
24	1225	83.0	243	15	US-10-058-270A-38	Sequence 38, Appli
25	1225	83.0	243	15	US-10-634-108-4	Sequence 4, Appli
26	1225	83.0	243	15	US-10-188-832-115	Sequence 175, App
27	1225	83.0	243	18	US-10-939-233-4	Sequence 4, Appli
28	1222	82.8	243	15	US-10-351-334-205	Sequence 205, Appli
29	1220	82.7	243	14	US-10-176-847-14	Sequence 14, Appli
30	1216	82.4	243	9	US-09-834-759-514	Sequence 514, App
31	1216	82.4	243	9	US-09-938-418-7	Sequence 7, Appli
32	1216	82.4	243	10	US-09-946-374-352	Sequence 352, App
33	1216	82.4	243	10	US-09-946-374-431	Sequence 431, App
34	1216	82.4	243	13	US-10-006-867-122	Sequence 122, App
35	1216	82.4	243	13	US-10-052-586-366	Sequence 366, App
36	1216	82.4	243	13	US-10-007-805-514	Sequence 514, App
37	1216	82.4	243	13	US-10-063-547-122	Sequence 122, App
38	1216	82.4	243	13	US-10-063-551-122	Sequence 122, App
39	1216	82.4	243	14	US-10-174-550-366	Sequence 366, App
40	1216	82.4	243	14	US-10-176-758-366	Sequence 366, App
41	1216	82.4	243	14	US-10-175-737-366	Sequence 366, App
42	1216	82.4	243	14	US-10-063-616-122	Sequence 122, App
43	1216	82.4	243	14	US-10-174-581-366	Sequence 366, App
44	1216	82.4	243	14	US-10-176-483-366	Sequence 366, App
45	1216	82.4	243	14	US-10-176-749-366	Sequence 366, App

#### ALIGNMENTS

RESULT 1									
US-10-045-992-5									
; Sequence 5, Application US//10045992									
; Publication No. US20020161211A1									
GENERAL INFORMATION:									
; APPLICANT: LINDNER, Volhard									
; APPLICANT: FRIESEL, Robert E.									
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN									
; FILE REFERENCE: 053689-5006-01									
; CURRENT APPLICATION NUMBER: US/10/045,992									
; CURRENT FILING DATE: 2001-10-19									
; PRIOR APPLICATION NUMBER: US 09/692,081									
; PRIOR FILING DATE: 2000-10-19									
; NUMBER OF SEQ ID NOS: 9									
; SOFTWARE: PatentIn version 3.1									
; SEQ ID NO 5									
; LENGTH: 277									
; TYPE: PRT									
; ORGANISM: Rattus sp.									
US-10-045-992-5									
Query Match									
Best Local Similarity 100.0%; Score 1476; DB 13; Length 277;									
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;									
QY	1	MRPAELGQTLSRAGLCRPL	1	MRPAELGQTLSRAGLCRPL	1	MRPAELGQTLSRAGLCRPL	1	MRPAELGQTLSRAGLCRPL	1
DB	1	MRPAELGQTLSRAGLCRPL	1	MRPAELGQTLSRAGLCRPL	1	MRPAELGQTLSRAGLCRPL	1	MRPAELGQTLSRAGLCRPL	1
QY	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61
DB	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61	SSASBNPKVKKALIRQREVDVLYNQCLOCPAVPGDGSFGANGIPGTGIGRDOF	61
QY	121	KKKKCKLRSPSESWTPVYKCCSSSLNYGIDGKIAECFTTKRNSALRVLPSSGLR	121	KKKKCKLRSPSESWTPVYKCCSSSLNYGIDGKIAECFTTKRNSALRVLPSSGLR	121	KKKKCKLRSPSESWTPVYKCCSSSLNYGIDGKIAECFTTKRNSALRVLPSSGLR	121	KKKKCKLRSPSESWTPVYKCCSSSLNYGIDGKIAECFTTKRNSALRVLPSSGLR	121

DB 121 KGEKECLRESFEESWTPTNYKOCWSLNYGIDLGKIACTFTKRSNSALRVLFSSGLR 180  
QY 181 LKCRNACCORWYFTTNGAECGSLPIEALITYLDQSPELNSTINIHRTSSVEGLCEGIGA 240  
DB 181 LKCRNACCORWYFTTNGAECGSLPIEALITYLDQSPELNSTINIHRTSSVEGLCEGIGA 240  
QY 241 GLVDVAIWGTCSDYKPGDASTGMNSVSRITIEELPK 277  
DB 241 GLVDVAIWGTCSDYKPGDASTGMNSVSRITIEELPK 277

## RESULT 2

US-10-634-108-5  
; Sequence 5, Application US/10634108  
; Publication No. US20040063176A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhart  
; APPLICANT: FRIESEL, Robert F.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL  
; FILE REFERENCE: 36-13 Lindner et al (200036.0013)  
; CURRENT APPLICATION NUMBER: US/10/634,108  
; PRIOR FILING DATE: 2003-08-04  
; PRIOR APPLICATION NUMBER: US/09/692,081  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-634-108-5

Query Match 100.0%; Score 1476; DB 15; Length 277;  
Best Local Similarity 100.0%; Pred. No. 3.4e-134;  
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPPAEIQTLSRAGLCRPLCLLCSQLPHTMHPQRAASPOLLLGLFVLLILLQLQSA 60  
DB 1 MPPAEIQTLSRAGLCRPLCLLCSQLPHTMHPQRAASPOLLLGLFVLLILLQLQSA 60  
QY 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
DB 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
QY 121 KGEKECLRESFEESWTPTNYKOCWSLNYGIDLGKIACTFTKRSNSALRVLFSSGLR 180  
DB 121 KGEKECLRESFEESWTPTNYKOCWSLNYGIDLGKIACTFTKRSNSALRVLFSSGLR 180  
QY 181 LKCRNACCORWYFTTNGAECGSLPIEALITYLDQSPELNSTINIHRTSSVEGLCEGIGA 240  
DB 181 LKCRNACCORWYFTTNGAECGSLPIEALITYLDQSPELNSTINIHRTSSVEGLCEGIGA 240  
QY 241 GLVDVAIWGTCSDYKPGDASTGMNSVSRITIEELPK 277  
DB 241 GLVDVAIWGTCSDYKPGDASTGMNSVSRITIEELPK 277

## RESULT 3

US-10-939-233-5  
; Sequence 5, Application US/10939233  
; Publication No. US20050147602A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhart  
; APPLICANT: FRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTNRC1, A NOVEL  
; FILE REFERENCE: 053689-5006-03  
; CURRENT APPLICATION NUMBER: US/10/939,233  
; PRIOR FILING DATE: 2004-09-10  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; PRIOR FILING DATE: 2000-10-19  
; PRIOR APPLICATION NUMBER: US 10/045,992

; PRIOR FILING DATE: 2001-10-19  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 5  
; LENGTH: 277  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-939-233-5

Query Match 100.0%; Score 1476; DB 18; Length 277;  
Best Local Similarity 100.0%; Pred. No. 3.4e-134;  
Matches 277; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MPPAEIQTLSRAGLCRPLCLLCSQLPHTMHPQRAASPOLLLGLFVLLILLQLQSA 60  
DB 1 MPPAEIQTLSRAGLCRPLCLLCSQLPHTMHPQRAASPOLLLGLFVLLILLQLQSA 60  
QY 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
DB 61 PSSASENPVKOKALIRQREVVDLYNGMCLQGPAGVPGRDSPGANGIPGTPIGRDGF 120  
QY 121 KGEKECLRESFEESWTPTNYKOCWSLNYGIDLGKIACTFTKRSNSALRVLFSSGLR 180  
DB 121 KGEKECLRESFEESWTPTNYKOCWSLNYGIDLGKIACTFTKRSNSALRVLFSSGLR 180  
QY 181 LKCRNACCORWYFTTNGAECGSLPIEALITYLDQSPELNSTINIHRTSSVEGLCEGIGA 240  
DB 181 LKCRNACCORWYFTTNGAECGSLPIEALITYLDQSPELNSTINIHRTSSVEGLCEGIGA 240  
QY 241 GLVDVAIWGTCSDYKPGDASTGMNSVSRITIEELPK 277  
DB 241 GLVDVAIWGTCSDYKPGDASTGMNSVSRITIEELPK 277

## RESULT 4

US-10-045-992-2  
; Sequence 2, Application US/10045992  
; Publication No. US20020161211A1  
; GENERAL INFORMATION:  
; APPLICANT: LINDNER, Volkhart  
; APPLICANT: FRIESEL, Robert E.  
; TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODELIN  
; FILE REFERENCE: 053689-5006-01  
; CURRENT APPLICATION NUMBER: US/10/045,992  
; PRIOR FILING DATE: 2001-10-19  
; PRIOR APPLICATION NUMBER: US 09/692,081  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 245  
; TYPE: PRT  
; ORGANISM: Rattus sp.  
US-10-045-992-2

Query Match 88.6%; Score 1307; DB 13; Length 245;  
Best Local Similarity 100.0%; Pred. No. 6.7e-118;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MPPQRAASPOLLLGLFVLLILLQLSAPSSASENPVKOKALIRQREVVDLYNGMCLQ 92  
DB 1 MPPQRAASPOLLLGLFVLLILLQLSAPSSASENPVKOKALIRQREVVDLYNGMCLQ 60  
QY 93 PAVGVRGDSPGANGIPGTPIGRDGFKGEKGECLRESFEESWTPTNYKOCWSLNYGI 152  
DB 61 PAVGVRGDSPGANGIPGTPIGRDGFKGEKGECLRESFEESWTPTNYKOCWSLNYGI 120  
QY 153 DKGIAECTFTKRSNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEALITYL 212  
DB 121 DKGIAECTFTKRSNSALRVLFSSGLRLKCRNACCORWYFTTNGAECGSLPIEALITYL 180  
QY 213 DQSPELNSTINIHRTSSVEGLCEGIGALVDVAIWGTCSDYKPGDASTGMNSVSRIT 272

Db 181 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSR111 240  
QY 273 EELPK 277  
241 EELPK 245

RESULT 5  
US-10-634-108-2

Sequence 2, Application US/10634108  
Publication No. US20040063176A1  
GENERAL INFORMATION:  
APPLICANT: LINDNER, Volkhard  
APPLICANT: PRIESEL, Robert F.  
TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO REMODEL  
FILE REFERENCE: 36-13 Lindner et al. (200036.0013)  
CURRENT APPLICATION NUMBER: US/10/634,108  
CURRENT FILING DATE: 2003-08-04  
PRIOR APPLICATION NUMBER: US/09/692,081  
PRIOR FILING DATE: 2000-10-19  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 245  
TYPE: PRT  
ORGANISM: Rattus sp.  
US-10-634-108-2

Query Match 88.6%; Score 1307; DB 15; Length 245;  
Best Local Similarity 100.0%; Pred. No. 6,7e-118;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRQREVVDLYNMCLOG 92  
Db 1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRQREVVDLYNMCLOG 60  
QY 93 PAVPGRDSPGANGIPGTGPIGRDGFKEGECCLRESPEESWTPTYKCCSSSLNYGI 152  
Db 61 PAVPGRDSPGANGIPGTGPIGRDGFKEGECCLRESPEESWTPTYKCCSSSLNYGI 120  
QY 153 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCRWYFTTNGACSGPLPIEAIYI 212  
Db 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCRWYFTTNGACSGPLPIEAIYI 180  
QY 213 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSR111 272  
Db 181 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSR111 240  
QY 273 EELPK 277  
Db 241 EELPK 245

RESULT 6

US-10-939-233-2  
Sequence 2, Application US/10939233  
Publication No. US20050147602A1  
GENERAL INFORMATION:  
APPLICANT: LINDNER, Volkhard  
APPLICANT: PRIESEL, Robert E.  
TITLE OF INVENTION: COMPOSITIONS, METHODS AND KITS RELATING TO CTNRC1, A NOVEL  
FILE REFERENCE: 053689-5006-03  
CURRENT APPLICATION NUMBER: US/10/939,233  
CURRENT FILING DATE: 2004-09-10  
PRIOR APPLICATION NUMBER: US 09/692,081  
PRIOR FILING DATE: 2000-10-19  
PRIOR APPLICATION NUMBER: US 10/045,992  
PRIOR FILING DATE: 2001-10-19  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 2  
LENGTH: 245

TYPE: PRT  
ORGANISM: Rattus sp.  
US-10-939-233-2

Query Match 88.6%; Score 1307; DB 18; Length 245;  
Best Local Similarity 100.0%; Pred. No. 6,7e-118;  
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRQREVVDLYNMCLOG 92  
Db 1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPVKOKALIRQREVVDLYNMCLOG 60  
QY 93 PAVPGRDSPGANGIPGTGPIGRDGFKEGECCLRESPEESWTPTYKCCSSSLNYGI 152  
Db 61 PAVPGRDSPGANGIPGTGPIGRDGFKEGECCLRESPEESWTPTYKCCSSSLNYGI 120  
QY 153 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCRWYFTTNGACSGPLPIEAIYI 212  
Db 121 DLGKIACFTTKRNSALRVLFSGSLRLKCRNACCRWYFTTNGACSGPLPIEAIYI 180  
QY 213 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSR111 272  
Db 181 DQSPPELNTSTINIHRTSSVEGLCEGIGAGLVDAIIVWGTCSDPYKGDASTGMNSVSR111 240  
QY 273 EELPK 277  
Db 241 EELPK 245

RESULT 7

US-10-097-340-131  
Sequence 131, Application US/10097340  
Publication No. US20030087250A1  
GENERAL INFORMATION:  
APPLICANT: John MONAHAN  
APPLICANT: Manjula GANNAVAPU  
APPLICANT: Sebastian HOERSCHE  
APPLICANT: Shubhangi KAMATKAR  
APPLICANT: Steve G. KOVATIS  
APPLICANT: Rachel E. MEYERS  
APPLICANT: Michael MORRISSEY  
APPLICANT: Peter OLANDT  
APPLICANT: Ami SEN  
APPLICANT: Peter VEIBY  
APPLICANT: Gordon B. MILLS  
APPLICANT: Robert C. BAST, Jr.  
APPLICANT: Karen LU  
APPLICANT: Rosemarie SCHMANDT  
APPLICANT: Xumei ZHAO  
APPLICANT: Karen GLATT  
TITLE OF INVENTION: Nucleic Acid Molecules and Proteins For The Identification,  
FILE REFERENCE: MRI-030  
CURRENT APPLICATION NUMBER: US/10/097,340  
CURRENT FILING DATE: 2002-03-14  
PRIOR APPLICATION NUMBER: 60/276,025  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/325,149  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/276,026  
PRIOR FILING DATE: 2001-03-14  
PRIOR APPLICATION NUMBER: 60/324,967  
PRIOR FILING DATE: 2001/09/26  
PRIOR APPLICATION NUMBER: 60/311,732  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: 60/325,102  
PRIOR FILING DATE: 2001-09-26  
PRIOR APPLICATION NUMBER: 60/323,580  
PRIOR FILING DATE: 2001-09-19  
NUMBER OF SEQ ID NOS: 363  
SOFTWARE: FaestSeq for Windows Version 4.0  
SEQ ID NO 131  
LENGTH: 278

[illegible]

```

RESULT 8
US-10-177-293-496
; Sequence 496, Application US/10177293
; Publication No. US20030124128A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Glatte, Karen
; APPLICANT: Zhao, Xumei
; APPLICANT: Ganavarpur, Manjula
; APPLICANT: Kamatekar, Shubhang
; APPLICANT: Mertens, Maureen
; APPLICANT: Myer, Vic
; APPLICANT: Wang, Youzhen
; APPLICANT: Xu, Yongyao
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Monahan, John
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Bast Jr., Robert C.
; APPLICANT: Hortobagyi, Gabriel N.
; APPLICANT: Pusztai, Lajos
; APPLICANT: Meric, Funda
; APPLICANT: Sahin, Aysegul
; APPLICANT: Mills, Gordon B.
; TITLE OF INVENTION: COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT,
; FILE REFERENCE: MRI-038
; PREVENTION, AND THERAPY OF BREAST CANCER
; CURRENT APPLICATION NUMBER: US/10/177,293
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US 60/299,887
; PRIOR FILING DATE: 2001-06-21
; PRIOR APPLICATION NUMBER: US 60/301,572
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: US 60/306,501
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/325,002
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/362,585
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/xxx,xxx
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 506
; SOFTWARE: PatSeq for Windows Version 4.0
; SEQ ID NO 496
; LENGTH: 278

```

	TYPE: PRT	ORGANISM: Homo sapiens	US-10-177-293-496
Query Match	84.2%	Score 1243.5; DB 14; Length 278;	
Best Local Similarity	86.8%	Pred. No. 1.le-11;	
Matches 243;	Conservative 6;	Mismatches 26;	Indels 5; Gaps 2
QY	1	MRPAAL--GQTSRAGLCRPCLLLCASQLPHTMHPOGRAASQQLLGLFLVLLILIQ	57
DB	1	MQPAASERGGADADHVPLGLLRQLQRAARQPMGRPOGPAASQQRARG--LLILLILIQ	58
QY	58	LSAPSSASENPKYKOKALIRQREYVDLNGMKLOSPAGPGDGSPPANGIGCTGPIGR	117
DB	59	LPAPSSAEIPIKQKOKAQLQREYVDLNGMKLOSPAGPGDGSPPANGIGCTGPIGR	118
QY	118	DGFKEKECECLRESEFESEMTPNYKQCSMSLNYGIDLCKIAECTPTKMSNLAHYLPSG	177
DB	119	DGFKEKECECLRESEFESEMTPNYKQCSMSLNYGIDLCKIACTPTKMSNLAHYLPSG	178
QY	178	SLRLKRNACCQRMVFTFNGABCSGPLIEALITYDQSPELINSTINIHRTSSVBGLCEG	237
DB	179	SLRLKRNACCQRMVFTFNGABCSGPLIEALITYDQSPPEMNSTINIHRTSSVBGLCEG	238
QY	238	IGAGLVDAIVAWGTGSDYPKGDASTGMSVSRILIIIEELPK	277
DB	239	IGAGLVDAIVAWGTGSDYPKGDASTGMSVSRILIIIEELPK	278

```

RESULT 9
US-10-301-822-77
; Sequence 77, Application US/10301822
; Publication No. US20030148410A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Berger, Allison
; APPLICANT: Guillemette, Tracy L.
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Schlegel, Robert
; APPLICANT: Monahan, John E.
; APPLICANT: Thibodeau, Stephen N.
; APPLICANT: Burgart, Lawrence J.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; FILE REFERENCE: MPM01-029P2RMM
; CURRENT APPLICATION NUMBER: US/10/301,822
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 60/339,971
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US 60/361,978
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: US 60/381,988
; PRIOR FILING DATE: 2002-05-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: PaedSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-10-301-822-77

Query Match      84.2%; Score 1243.5; DB 14; Length 278;
Best Local Similarity 86.8%; Pred. No. 1.le-111;
Matches 243; Conservative 6; Mismatches 26; Indels 5; Gaps 2;

OY      1 KRPAAL--GQTSRAGLCRPLCLLCASQLPTMHPOGRAASPOLLIGFLVLLDLQ 57
          |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      1 MQPAASBRGGADHDHVLGLILRLQLFAARQPGAMRPOGPAASFQRURG--LTLILLDQ 58

OY      58 LSAPSSAEENKRYOKALLIRREVVDLYNGMCLOGPAGVPERDDSPGANGIPIGPRIGR 117
          ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
DB      59 LPAPSASEIRPKQKQAKLRREVVLDLYNGMCLGSPAGVPERDDSPGANIGIPGTGLPR 118

```

Qy	118	DGFEGKEGCEIARSPFEESWT	PVNYKQCSWSLWYDGLGACETTKMRSSALRYL	FSG	177
Db	119	DGFEGKEGCEIARSPFEESWT	PVNYKQCSWSLWYDGLGACETTKMRSSALRYL	FSG	178
Qy	178	SLILKCRNACCCMYWTFNPGAECSGLP	LEAIIYYIDQSPELNNTINHTRSSV	CEG	237
Db	179	SLILKCRNACCCMYWTFNPGAECSGLP	LEAIIYYIDQSPENNTINHTRSSV	CEG	238
Qy	238	IGAGLYDVVAIWTGSCDYPKGDASTGMSVSR	IIIIIEELPK	277	
Db	239	IGAGLYDVVAIWTGSCDYPKGDASTGMSVSR	IIIIIEELPK	278	

```

RESULT 10
US-10-961-139-2
; Sequence 2, Application US/10961139
; Publication No. US2005015313A1
; GENERAL INFORMATION:
; APPLICANT: Endge, Wilson O.
; APPLICANT: Ford, Donna
; APPLICANT: Gannavarapu, Manjula
; APPLICANT: Glatt, Karen
; APPLICANT: Hoersch, Sebastian
; APPLICANT: Kamatkar, Shubhangi
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; APPLICANT: Xu, Yong Yao
; APPLICANT: Zhao, Xumei
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES AND PROTEINS FOR
; TITLE OF INVENTION: THE IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: OVARIAN CANCER
; FILE REFERENCE: WRI-065
; CURRENT APPLICATION NUMBER: US/10/961.139
; CURRENT FILING DATE: 2004-10-07
; PRIOR APPLICATION NUMBER: 60/509,171
; PRIOR FILING DATE: 2003-10-07
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapien
; US-10-961-139-2

```

[illegible]

```

? Publication No. US20030073144A1
? GENERAL INFORMATION:
? APPLICANT: Benson, Darin R.
? APPLICANT: Kalos, Michael D.
? APPLICANT: Lodes, Michael J.
? APPLICANT: Persing, David H.
? APPLICANT: Hepler, William T.
? APPLICANT: Jiaang, Yugu
? TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
? TITLE OF INVENTION: AND DIAGNOSIS OF PANCREATIC CANCER
? FILE REFERENCE: 210121.566
? CURRENT APPLICATION NUMBER: US/10/060.036
? CURRENT FILING DATE: 2002-01-30
? NUMBER OF SEQ ID NOS: 4560
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 4551
?
? LENGTH: 278
?
? TYPE: PRT
?
? ORGANISM: Homo sapiens
?
? US-10-060-036-4551

```

[illegible]

```

      RESULT 12
US-10-296-115-1261
; Sequence 1261, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
APPLICANT : Hyseq Inc
TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
FILE REFERENCE: 784PCF
CURRENT APPLICATION NUMBER: US/10/296.115
PRIOR FILING DATE: 2002-11-18
PRIOR APPLICATION NUMBER: US09/488,725
PRIOR FILING DATE: 2000-01-21
PRIOR APPLICATION NUMBER: US09/552,317
PRIOR FILING DATE: 2000-04-25
NUMBER OF SEQ ID NOS: 1478
SEQ ID NO 1261
LENGTH: 278
TYPE: PRt
ORGANISM: Homo sapiens
US-10-296-115-1261

      84.0%; Score 1239.5; DB 15; Length 278;
Query Match          Pred.No.2.7e-11;
Best Local Similarity 90.9%;
Matches    239;   Mismatches    14; Indels     5; Gaps     2

```

```
Db 21 GLLR---LQLRAARQPGAMRPOGPAAAPQRLRG--LTLTLTLTLQLPAPSSASEIPKQKXA 75
Qy 75 LIRQREVVDLYNGMCLQGPAGVPGRGSGFANGIPCTPGIPGRDQKGEKGECLRSFEE 134
Db 76 QLRQREVVDLYNGMCLQGPAGVPGRGSGFANGIPCTPGIPGRDQKGEKGECLRSFEE 135
Qy 135 SWTPNKKQCSWSLNYGIDLGKIAECTPTKMRNSALRYLFSGSLTKCRNACCOMWYFT 194
Db 136 SWTPNKKQCSWSLNYGIDLGKIAECTPTKMRNSALRYLFSGSLTKCRNACCOMWYFT 195
Qy 195 FNGAECGSPLEPIEAIYYLDQSPBELNSTINIHRTSSVEGLCEGIGALVDVAIWGTCS 254
Db 196 FNGAECGSPLEPIEAIYYLDQSPBELNSTINIHRTSSVEGLCEGIGALVDVAIWGTCS 255
Qy 255 YPKGASTGMSVSRITIIIEELPK 277
Db 256 YPKGASTGMSVSRITIIIEELPK 278
```

## RESULT 13

```
US-09-834-759-515
; Sequence 515, Application US/09834759
; Publication No. US20020085998A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguin
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C9
; CURRENT APPLICATION NUMBER: US/09/834, 759
; NUMBER OF SEQ ID NOS: 547
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-834-759-515
```

Query Match 83.6%; Score 1234.5; DB 9; Length 278;

Best Local Similarity 86.4%; Pred. No. 8.1e-111; Mismatches 27; Indels 5; Gaps 2;

Matches 242; Conservative 6;

```
Qy 1 MRPAEL---GOTLSRAGLCRPLCLLCAQOLPHTMHPOGRAAPQLLGLFLVLLLLIQ 57
Db 1 MQPAAASERGADADHVPFLGLRLQLRAARQPGAMRPOGPAAAPQRLRG--LTLTLTL 58
Qy 58 LSAPSSASBNPKYKQKALIRQREVVDLYNGMCLQGPAGVPGRGSGFANGIPCTPGIPGR 117
Db 59 LPAPSSASEIPKQKQALRQREVVDLYNGMCLQGPAGVPGRGSGFANGIPCTPGIPGR 118
Qy 118 DQFKGEKGECLRSFEEESWTPNKKQCSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 177
Db 119 DQFKGEKGECLRSFEEESWTPNKKQCSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 178
Qy 178 SLRLKCRNACCOMWYFTFNGAECGSPLEPIEAIYYLDQSPBELNSTINIHRTSSVEGLCEG 237
Db 179 SLRLKCRNACCOMWYFTFNGAECGSPLEPIEAIYYLDQSPBELNSTINIHRTSSVEGLCEG 238
Qy 238 IGAGLVDAVAIWGTCSDPYKGDASTGMNSVSRITIIIEELPK 277
Db 239 IGAGLVDAVAIWGTCSDPYKGDASTGMNSVSRITIIIEELPK 278
```

## RESULT 14

```
US-10-007-805-515
; Sequence 515, Application US/10007805
; Publication No. US20020150581A1
```

GENERAL INFORMATION:

```
; APPLICANT: Jiang, Yuguin
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiangchun
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Durham, Margareta
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C10
; CURRENT APPLICATION NUMBER: US/10/007,805
; NUMBER OF SEQ ID NOS: 593
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-007-805-515
```

Query Match 83.6%; Score 1234.5; DB 13; Length 278;

Best Local Similarity 86.4%; Pred. No. 8.1e-111; Mismatches 27; Indels 5; Gaps 2;

Matches 242; Conservative 6;

```
Qy 1 MRPAEL---GOTLSRAGLCRPLCLLCAQOLPHTMHPOGRAAPQLLGLFLVLLLLIQ 57
Db 1 MQPAAASERGADADHVPFLGLRLQLRAARQPGAMRPOGPAAAPQRLRG--LTLTLTL 58
Qy 58 LSAPSSASBNPKYKQKALIRQREVVDLYNGMCLQGPAGVPGRGSGFANGIPCTPGIPGR 117
Db 59 LPAPSSASEIPKQKQALRQREVVDLYNGMCLQGPAGVPGRGSGFANGIPCTPGIPGR 118
Qy 118 DQFKGEKGECLRSFEEESWTPNKKQCSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 177
Db 119 DQFKGEKGECLRSFEEESWTPNKKQCSWSLNYGIDLGKIAECTPTKMRNSALRYLFSG 178
Qy 178 SLRLKCRNACCOMWYFTFNGAECGSPLEPIEAIYYLDQSPBELNSTINIHRTSSVEGLCEG 237
Db 179 SLRLKCRNACCOMWYFTFNGAECGSPLEPIEAIYYLDQSPBELNSTINIHRTSSVEGLCEG 238
Qy 238 IGAGLVDAVAIWGTCSDPYKGDASTGMNSVSRITIIIEELPK 277
Db 239 IGAGLVDAVAIWGTCSDPYKGDASTGMNSVSRITIIIEELPK 278
```

## RESULT 15

```
US-10-076-622-515
; Sequence 515, Application US/10076622
; Publication No. US20030023036A1
; GENERAL INFORMATION:
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Persing, David H.
; APPLICANT: Sleath, Paul R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.470C11
; CURRENT APPLICATION NUMBER: US/10/076,622
; NUMBER OF SEQ ID NOS: 627
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 515
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-076-622-515
```

Query Match

Best Local Similarity 86.4%; Score 1234.5; DB 14; Length 278;

Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

Qy	1	MRPAEL--GQTLNRAGLCRPICLLICASQLPHITMHPGGRASPQULLGLFVLLHLLQ	57
Db	1	MDPAASERGGADADHVPFLGLRLQLRAARQPGMRPGPAAAPQRLRG--LILLHLLQ	58
Qy	58	LSAPSSASENPVKYKALIRQREVVDLYNGMCLQGPAGVPGRDGSPGANGIPGTPIGR	117
Db	59	LPAPSSASEIPKQKQKQLQREVVDLYNGMCLQGPAGVPGRDGSPGANVIPGTPIGR	118
Qy	118	DGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDLGLIAECTFTKMRSNSALRYLFSG	177
Db	119	DGFKEGKECLRESFEESWTPTYKQCSWSSLYNGIDLGLIAECTFTKMRSNSALRYLFSG	178
Qy	178	SLRLKCRNACCCORWYFTFENGACSGPLPIAIIYLDQSPELNSTINIHRTSSVEGLCEG	237
Db	179	SLRLKCRNACCCORWYFTFENGACSGPLPIAIIYLDQSPENNSTINIHRTSSVEGLCEG	238
Qy	238	IGAGLVDAIIVGTCSDPYKGDASTGWNVSRIIIEELPK	277
Db	239	IGAGLVDAIIVGTCSDPYKGDASTGWNVSRIIIEELPK	278

Search completed: August 4, 2005, 04:45:28  
 Job time : 99.1268 secs

This Page Blank (uspto)



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 4, 2005, 04:04:29 ; Search time 106.817 Seconds

(without alignments)  
1002.955 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAAELGGLTSLRAGLCRPL.....GDASTGWSVSRRIIELEPK 277

## Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1476	100.0	277	5	ABBB0979
2	1476	100.0	277	5	ABG75759
3	1307	88.6	245	5	ABBB0977
4	1307	88.6	245	6	ABG75757
5	1287.5	87.2	244	8	ADOC28703
6	1243.5	84.2	278	3	AAAB08856
7	1243.5	84.2	278	5	ABG96338
8	1243.5	84.2	278	5	ABR47627
9	1240.5	84.0	278	5	ABP68631
10	1239.5	84.0	278	4	AAAM25746
11	1234.5	83.6	278	5	ABP68637
12	1234.5	83.6	278	5	ABG78939
13	1234.5	83.6	278	6	ABJ37762
14	1234.5	83.6	278	7	ADJ93177
15	1225	83.0	243	4	ABAB6667
16	1225	83.0	243	5	ABG96340
17	1225	83.0	243	5	ABJ05554
18	1225	83.0	243	5	ABBB0978
19	1225	83.0	243	6	ABR48526
20	1225	83.0	243	6	ABR48227
21	1225	83.0	243	6	ABG75758
22	1225	83.0	243	6	ABU56607
23	1225	83.0	243	7	ADNB0510
24	1225	83.0	243	7	ADNB9855
25	1225	83.0	243	7	ADNB38732

26	1225	83.0	243	8	ADJ70254	ADJ70254	IBFL301	P
27	1222	82.8	243	3	AA991529	AA991529	Human	sec
28	1222	82.8	243	8	ADJ71601	ADJ71601	Novel	hum
29	1220	82.7	243	6	ABJ37031	ABJ37031	Human	bre
30	1216	82.4	243	3	AA994462	AA994462	Human	PRO
31	1216	82.4	243	3	AA994441	AA994441	Human	PRO
32	1216	82.4	243	4	AB66190	AB66190	Protein	O
33	1216	82.4	243	4	AB66221	AB66221	Protein	O
34	1216	82.4	243	4	AAU29206	AAU29206	Human	PRO
35	1216	82.4	243	4	AB87586	AB87586	Human	PRO
36	1216	82.4	243	5	ABG95911	ABG95911	Human	sec
37	1216	82.4	243	5	ABBB4939	ABBB4939	Human	PRO
38	1216	82.4	243	5	AAE20462	AAE20462	Human	tum
39	1216	82.4	243	5	ABR95545	ABR95545	Human	ang
40	1216	82.4	243	5	ABP68636	ABP68636	Human	pan
41	1216	82.4	243	5	ABG78938	ABG78938	Human	bre
42	1216	82.4	243	6	ABU58582	ABU58582	Human	PRO
43	1216	82.4	243	6	ABU88130	ABU88130	Novel	hum
44	1216	82.4	243	6	ABU84445	ABU84445	Human	sec
45	1216	82.4	243	6	ABR66319	ABR66319	Human	sec

## ALIGNMENTS

RESULT 1  
ABBB0979  
ID ABBB0979 standard; protein; 277 AA.

AC ABBB0979;

DT 21-OCT-2002 (first entry)

DE Long form of rat REMODELIN polypeptide.

KM REMODELIN; vulnery; vasotropic; cytosstatic; osteopathic; collagen;

KW gene therapy; bone; rat.

XX Rattus sp.

XX WO200242487-A2.

XX 30-MAY-2002.

XX 19-OCT-2001; 2001WO-US050940.

XX 19-OCT-2000; 2000US-00692081.

XX (MAIN-) MAINE MEDICAL CENT RES INST.

XX Linder V, Friesel R;

XX WPI; 2002-590472/63.

XX N-PSDB; ABR6480.

PT New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
negative remodeling, fibrosis, collagen disease, and bone disease in a  
mammal.

XX Claim 5; Fig 4C; 187pp; English.

CC The invention relates to mammalian REMODELIN polypeptides and encoding  
CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
CC negative remodeling or fibrosis, bone disease such as osteogenesis  
CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
CC bullosa (DEB) and Bethlem myopathy in a mammal. The polynucleotides are  
CC used to affect cellular gene expression in a mammal, where the cellular  
CC gene is from transforming growth factor (TGF)-beta1, collagen fibrillar  
CC osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein  
CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
CC comprising antisense REMODELIN sequences are useful for treating diseases  
CC mediated by abnormal expression of a REMODELIN molecule in a human such  
CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or



XX 21-OCT-2002 (first entry)  
 DT Rat REMODELIN polypeptide.  
 XX REMODELIN; vulnery; vasotropic; cytostatic; osteopathic; collagen;  
 KM gene therapy; bone; rat.  
 XX Rattus sp.  
 OS  
 PN W0200242487-A2.  
 XX 30-MAY-2002.  
 PD  
 XX 19-OCT-2001; 2001MO-US050940.  
 PF  
 XX 19-OCT-2000; 2000US-00692081.  
 PR  
 XX (MAIN-) MAINE MEDICAL CENT RES INST.  
 PA  
 PI Linder V, Friesel R;  
 DR WPI; 2002-590472/63.  
 XX N-PSDB; ABN86480.  
 DR  
 XX New mammalian REMODELIN polypeptide for diagnosing arterial restenosis,  
 PT negative remodeling, fibrosis, collagen disease, and bone disease in a  
 PT mammal.  
 PT  
 XX Claim 5; Fig 4B; 187pp; English.  
 PS  
 XX The invention relates to mammalian REMODELIN polypeptides and encoding  
 CC polynucleotides. REMODELIN is used to diagnose arterial restenosis,  
 CC negative remodeling or fibrosis, bone disease such as osteogenesis  
 CC imperfecta (OI), collagen disease such as OI, dystrophic epidermolysis  
 CC bullosa (DEB) and Bechlem myopathy in a mammal. The polynucleotides are  
 CC used to affect cellular gene expression in a mammal, where the cellular  
 CC gene is from transforming growth factor (TGF)-beta1, collagen Itrifalpa,  
 CC osteopontin, biglycan, alkaline phosphatase or bone morphogenic protein  
 CC 4. The expression of osteopontin is dependent on Cbfa1. Compositions  
 CC comprising antisense REMODELIN sequences are useful for treating diseases  
 CC mediated by abnormal expression of a REMODELIN molecule in a human such  
 CC as impaired wound healing, fibrosis of an organ, ectopic ossification, or  
 CC hypertrophic scar formation. REMODELIN is useful in the development of  
 CC bone during mammalian embryogenesis, proliferation and/or migration,  
 CC cellular signaling, adventitial fibrosis, negative remodeling and arterial  
 CC restenosis, smooth muscle cell proliferation, and arterial remodeling and  
 CC useful in gene therapy. REMODELIN is useful for elucidating the function  
 CC of REMODELIN molecules in a cell, to identify a compound that affects  
 CC REMODELIN expression and/or TGF-beta signaling, as a potential  
 CC therapeutic drug candidate for arterial restenosis, anti-cancer therapy,  
 CC to promote or inhibit wound healing, to inhibit scar tissue or keloid  
 CC formation, to promote bone fracture healing, and to increase or decrease  
 CC expression of REMODELIN in mammal. It is also useful for producing  
 CC recombinant cell and transgenic non-human mammals which are useful tools  
 CC for the study of REMODELIN action, for identifying novel diagnostics and  
 CC therapeutics for treatment, and for elucidating the cellular roles of  
 CC REMODELIN. The present sequence represents a rat REMODELIN polypeptide  
 XX  
 XX  
 SQ Sequence 245 AA;  
 Query Match 88.6%; Score 1307; DB 5; Length 245;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-122;  
 Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 33 MHPOGAAAPQLLIGFLVLLILQTSAPSSASNPVKOKALIRPQEVVDLYNGKLCQG 92  
 DB 1 MHPOGAAAPQLLIGFLVLLILQTSAPSSASNPVKOKALIRPQEVVDLYNGKLCQG 60  
 QY 93 PACVPGRDSPGANGIPGTPGIPGRDGFKEGKECELRSEFESWTNYYKQSSWSSLYNGI 152  
 DB 61 PACVPGRDSPGANGIPGTPGIPGRDGFKEGKECELRSEFESWTNYYKQSSWSSLYNGI 120

QY 153 DLGKIAECTPTKMRNSALRYLFGSGLRLKCRNACCCORRYFTFNGACGGPLFEATLYL 212  
 DB 121 DLGKIAECTPTKMRNSALRYLFGSGLRLKCRNACCCORRYFTFNGACGGPLFEATLYL 180  
 QY 213 DQGSPELNSTINIRHTSSVGLCEGIGAGLVDAVIAWGTCSDDPKDASTGMNSVSR111 272  
 DB 181 DQGSPELNSTINIRHTSSVGLCEGIGAGLVDAVIAWGTCSDDPKDASTGMNSVSR111 240  
 QY 273 EELPK 277  
 DB 241 EELPK 245  
 RESULT 4  
 ABG75757  
 ID ABG75757 standard; protein; 245 AA.  
 XX  
 AC ABG75757;  
 XX  
 DT 28-APR-2003 (first entry)  
 XX  
 XX Rat REMODELIN protein.  
 DE  
 XX Rat; adventitia-inducible; REMODELIN; REMODEL;  
 KM transforming growth factor beta; adventitia; vascular remodeling;  
 KM restenosis; vascular injury; antisense therapy; TGF-beta signaling;  
 KM TGF-beta receptor type II; arterial remodeling; bone formation;  
 KM cartilage formation; osteogenesis imperfecta; Bechlem myopathy;  
 KM dystrophic epidermolysis bullosa; negative remodeling; wound healing;  
 KM arterial stenosis; fibrosis; calcification; transplant; heart valve  
 KM heart valve transplant; osteopathic; antiarteriosclerotic; vulnery.  
 XX  
 XX Rattus sp.  
 OS  
 XX US2002161211-A1.  
 PN  
 XX 31-OCT-2002.  
 PD  
 XX 19-OCT-2001; 2001US-00045992.  
 PF  
 XX 19-OCT-2000; 2000US-00692081.  
 PR  
 XX (LIND/) LINDER V.  
 PA (FRIE/) FRIESEL R E.  
 PI Linder V, Friesel R;  
 DR WPI; 2003-238238/23.  
 XX N-PSDB; ABX11340.  
 DR  
 XX New isolated REMODELIN nucleic acid and polypeptide, useful for mediating  
 PT arterial remodeling, formation of bone and cartilage, and the diagnosis  
 PT and treatment of disorders associated with aberrant expression of  
 PT REMODELIN.  
 PT  
 XX Claim 3; Fig 4B; 81pp; English.  
 PS  
 XX The invention discloses an isolated nucleic acid encoding a mammalian  
 CC adventitia-inducible bone expressed molecule called REMODELIN (or REMODEL  
 CC and/or adventitia induced bone expressed molecule, AIBR). REMODELIN  
 CC expression was induced by transforming growth factor beta (TGF-beta),  
 CC which is important because proliferative events occurring in the  
 CC adventitia contribute to vascular remodeling and restenosis in the  
 CC to vascular injury and TGF-beta has been shown to be a factor involved in  
 CC this. Also disclosed is an antibody raised against REMODELIN and methods  
 CC for treating a disease mediated by abnormal expression of a REMODELIN in  
 CC a human, for identifying a compound that affects or reduces expression of  
 CC REMODELIN in a cell (e.g. antisense therapy), for identifying a compound  
 CC that affects TGF-beta signaling and for increasing or reducing REMODELIN  
 CC expression in a mammal, comprising administering a REMODELIN expression  
 CC increasing or reducing TGF-beta to the mammal, thereby increasing  
 CC REMODELIN expression or inhibiting signaling via the TGF-beta receptor

CC type II and reducing expression of REMODELIN in the mammal. The methods  
 CC and compositions of the present invention are useful for mediating  
 CC arterial remodeling, formation of bone and cartilage and the diagnosis  
 CC and treatment of disorders associated with aberrant expression of  
 CC REMODELIN, such as osteogenesis imperfecta, dystrophic epidermolysis  
 CC bullosa, Bethlem myopathy, negative remodeling, wound healing, arterial  
 CC stenosis, vessel injury, fibrosis and calcification of a transplant,  
 CC preferably a heart valve transplant. The sequence presented is the rat  
 CC REMODELIN protein

XX Sequence 245 AA;

Query Match 88.6%; Score 1307; DB 6; Length 245;

Best Local Similarity 100.0%; Pred. No. 1.9e-122; Indels 0; Gaps 0;

Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 33 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 92  
 DB 1 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 60

QY 93 PAGVPRDGRSPGANGIPGTPTGIPGRDGFKEGEGELRESFESWTNPKQCSWSSLYNGI 152

DB 61 PAGVPRDGRSPGANGIPGTPTGIPGRDGFKEGEGELRESFESWTNPKQCSWSSLYNGI 120

QY 153 DLGKIACEFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPIPIAIIYL 212

DB 121 DLGKIACEFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPIPIAIIYL 180

QY 213 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSRIII 272

DB 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSRIII 240

QY 273 EELPK 277

DB 241 EELPK 245

RESULT 5

ADO28703

ID ADO28703 standard; protein; 244 AA.

AC ADO28703;

DT 12-AUG-2004 (first entry)

XX HA4, SEQ ID 2.

XX Osteopathic; Cytostatic; Antiinflammatory; Antiarthritic; Antiheumatic;

KM osteoblast and chondrocyte-specific small secreted peptide; HA4;

KM bone degenerative disease; osteoporosis;

KM glucocorticoid induced osteoporosis; Paget's disease;

KM periodontal disease; tooth loss; bone fracture; rheumatoid arthritis;

KM periprosthetic osteolysis; osteogenesis imperfecta;

KM metastatic bone disease; hypercalcemia of malignancy.

XX Unidentified.

OS MO2004041205-A2.

XX 21-MAY-2004.

XX 04-NOV-2003; 2003MO-US035139.

XX 04-NOV-2002; 2002US-0423690P.

XX (TEXA ) UNIV TEXAS SYSTEM.

XX De Crombrughe B, Akiyama H;

PT peptide referred as HA4 polypeptide, useful for diagnosing or treating  
 PT bone degenerative disease, and for screening drugs that alter bone  
 PT formation.  
 PS Claim 1; SEQ ID NO 2; 103pp; English.  
 XX The present invention relates to an isolated osteoblast- and chondrocyte-  
 CC specific small secreted peptide named HA4 polypeptide (I; ADO28703) and  
 CC its coding sequence (II; ADO28702). (I) or (II) are useful for diagnosing  
 CC or creating a bone degenerative disease in a subject such as  
 CC osteoporosis, glucocorticoid induced osteoporosis, Paget's disease,  
 CC abnormally increased bone turnover, periodontal disease, tooth loss, bone  
 CC fractures, rheumatoid arthritis, periprosthetic osteolysis, osteogenesis  
 CC imperfecta, metastatic bone disease and hypercalcemia of malignancy. (II)  
 CC is also useful for identifying a subject at risk of or suffering from a  
 CC bone degenerative disease.

XX Sequence 244 AA;

Query Match 87.2%; Score 1287.5; DB 8; Length 244;

Best Local Similarity 99.2%; Pred. No. 1.7e-120; Indels 1; Gaps 1;

Matches 243; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 33 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 92

DB 1 MHPOGRAASPOLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCLQG 59

QY 93 PAGVPRDGRSPGANGIPGTPTGIPGRDGFKEGEGELRESFESWTNPKQCSWSSLYNGI 152

DB 60 PAGVPRDGRSPGANGIPGTPTGIPGRDGFKEGEGELRESFESWTNPKQCSWSSLYNGI 119

QY 153 DLGKIACEFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPIPIAIIYL 212

DB 120 DLGKIACEFTTKRNSALRVLPFSGSLRLKCRNACCORWYFTFNAGCSGPIPIAIIYL 179

QY 213 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSRIII 272

DB 180 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAIIVGTCSDYPKGDASTGMSVSRIII 239

QY 273 EELPK 277

DB 240 EELPK 244

RESULT 6

AAB08856

ID AAB08856 standard; protein; 278 AA.

AC AAB08856;

DT 15-JAN-2001 (first entry)

XX Amino acid sequence of a human secretory protein.

XX Human; secretory protein; HSECP; cancer; gastrointestinal disorder;

KM inflammation; cardiovascular disorder; neurological disorder.

XX Homo sapiens.

OS Homo sapiens.

XX Key

XX Peptide

XX Modified-site

XX Modified-site

XX Modified-site

XX Modified-site



```

Db      1 MOPAAASERGGADADHVPLGLRLQLRAAPQGMRRPOGPAAAPQRLRG--LTLTLTLQ 58
QY      58 LSAPSSASENPVKOKALIRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 117
Db      59 LPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 118
QY      118 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTFTKRRSNSALRVLPFG 177
Db      119 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTFTKRRSNSALRVLPFG 178
QY      178 SLRLKCRNAACCGRWYFTFNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEG 237
Db      179 SLRLKCRNAACCGRWYFTFNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEG 238
QY      238 IGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRILIEELPK 277
Db      239 IGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRILIEELPK 278

RESULT 8
ABR47627
ID      ABR47627 standard; protein; 278 AA.
XX
AC      ABR47627;
XX
DT      12-JUN-2003 (first entry)
DE      Breast cancer associated protein sequence SEQ ID NO:496.
XX
KW      Human; breast cancer; cytostatic; gene therapy.
XX
OS      Homo sapiens.
XX
PN      MO2003004969-A2.
XX
PD      16-JAN-2003.
XX
PF      21-JUN-2002; 2002MO-US019669.
XX
PR      21-JUN-2001; 2001US-0299887P.
PR      27-JUN-2001; 2001US-0301572P.
PR      18-JUL-2001; 2001US-0306501P.
PR      25-SEP-2001; 2001US-0325002P.
PR      05-MAR-2002; 2002US-0362585P.
PR      14-MAY-2002; 2002US-0380391P.
XX
PA      (MILL-) MILLENIUM PHARM INC.
XX
PI      Lillie J, Gannavarapu M, Glatc K, Hoeresh S, Kamatkar S;
PI      Mertens M, Monahan JE, Myer V, Wang Y, Xu Y, Zhao X, Meyers RE;
PI      Bast RC, Hortobagyi GN, Pusztai L, Meric F, Sahin A, Mills GB;
XX
DR      WPI; 2003-210361/20.
XX
DR      N-PSDB; ACC50329.
XX
PT      Breast cancer diagnosis or treatment by comparing the level of expression
PT      of a marker in a patient sample with that in the control non-breast
PT      cancer sample.
XX
PS      Claim 1; SEQ ID NO 496; 128bp; English.
XX
CC      The present invention describes a method for assessing whether a patient
CC      is afflicted with breast cancer. The method comprises comparing a patient
CC      of expression of a marker (gene/polypeptide see ACC50076 to ACC50334 and
CC      ABR47386 to ABR47632) in a patient sample and the normal level of
CC      expression of the marker in a control non-breast cancer sample, where a
CC      significant increase in the level of expression of the marker in the
CC      patient sample and the normal level is an indication that the patient is
CC      afflicted with breast cancer. The breast cancer associated sequences from
CC      the present invention have cytostatic activities and can be used in gene
CC      therapy. The method is useful for diagnosing and treating breast cancer.
CC      N.B. The sequence data for this patent did not form part of the printed
CC      specification, but was obtained in electronic format directly from WIPO
```

```

CC      at ftp.wipo.int/pub/published_pct_sequences
XX
SQ      Sequence 278 AA;
XX
Query Match      84.2%; Score 1243.5; DB 6; Length 278;
Best Local Similarity 86.8%; Pred. No. 5,5e-116;
Matches 243; Conservative
XX
QY      1 MRPAAEL---GQTLRAAGCRPLCLLCAQQLPHTHMQGAAAPQLLGLFLVYLLTLQ 57
Db      1 MOPAAASERGGADADHVPLGLRLQLRAAPQGMRRPOGPAAAPQRLRG--LTLTLTLQ 58
QY      58 LSAPSSASENPVKOKALIRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 117
Db      59 LPAPSSASEIPKQKQKQALRQREVVDLYNGMCLQSPAGVPGRGSGANGIPGTGPIGR 118
QY      118 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTFTKRRSNSALRVLPFG 177
Db      119 DGFKEGKEGCLRESFEESWTPNYKQCSWSLNYGIDLKIAECTFTKRRSNSALRVLPFG 178
QY      178 SLRLKCRNAACCGRWYFTFNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEG 237
Db      179 SLRLKCRNAACCGRWYFTFNGAECGSLPIEALIIYDQSPPELNTSTINIHRTSSVEGLCEG 238
QY      238 IGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRILIEELPK 277
Db      239 IGAGLVDAIIVWGTCSDPYKGDASTGMNSVSRILIEELPK 278

RESULT 9
ABP66631
ID      ABP66631 standard; protein; 278 AA.
XX
AC      ABP66631;
XX
DT      14-JAN-2003 (first entry)
DE      Human pancreatic cancer expressed protein SEQ ID NO 4551.
XX
KW      Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
KW      cytostatic; tumour.
XX
OS      Homo sapiens.
XX
PN      MO200260317-A2.
XX
PD      08-AUG-2002.
XX
PF      30-JAN-2002; 2002MO-US002781.
XX
PR      30-JAN-2001; 2001US-0265305P.
PR      31-JAN-2001; 2001US-0265682P.
PR      09-FEB-2001; 2001US-0267568P.
PR      21-MAR-2001; 2001US-0276651P.
PR      28-APR-2001; 2001US-0287112P.
PR      16-MAY-2001; 2001US-0291631P.
PR      12-JUL-2001; 2001US-0305484P.
PR      20-AUG-2001; 2001US-0313999P.
PR      27-NOV-2001; 2001US-0333626P.
XX
PA      (CORI-) CORIXA CORP.
XX
PI      Benson DR, Kloe MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
DR      WPI; 2002-627435/67.
XX
DR      N-PSDB; ABV99139.
XX
PT      New isolated polynucleotide and pancreatic tumor polypeptides, useful for
PT      diagnosing, preventing and/or treating cancer, particularly pancreatic
PT      cancer.
XX
PS      Claim 2; SEQ ID NO 4551; 300bp + Sequence Listing; English.
XX
XX
```

CC The invention relates to an isolated polynucleotide (I) comprising: (a)  
CC any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)  
CC complements of (a); (c) sequences consisting of at least 20 contiguous  
CC residues of (a); (d) sequences that hybridize to (a), under moderately  
CC stringent conditions; (e) sequences having at least 75% or 90% identity  
CC to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-  
CC ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer  
CC in a patient and compositions comprising polypeptides, polynucleotides,  
CC antibodies, fusion proteins, T cell populations and antigen presenting  
CC cells expressing the polypeptide are useful in treating pancreatic cancer  
CC and stimulating an immune response. The polynucleotides can be used as  
CC probes or primers for nucleic acid hybridisation. In the design and  
CC preparation of ribozyme molecules for inhibiting expression of the tumour  
CC polypeptides and proteins in the tumour cells, in vaccines and for gene  
CC therapy. Note: The sequence data for this patent did not form part of the  
CC printed specification, but was obtained in electronic format directly  
CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 278 AA:

Query Match 84.0%; Score 1240.5; DB 5; Length 278;  
Best Local Similarity 86.4%; Pred. No. 1.1e-115;  
Matches 242; Conservative 7; Mismatches 26; Indels 5; Gaps 2;

QY 1 MRPAEL---GQTLNRAGLCRPICLLCASQLPHTMHPOGRAASPOLLGLPLVLLLLQ 57  
DB 1 MGRPAASERGGADADHVPILGLRLQLRAKRGAMRPGAPASPORLRG--LLLLLLQ 58  
QY 58 LSAPSSASSENPKYKOKARLQREVDLYNGMCLQGPAGVGRDGPANGIPETPGIPGR 117  
DB 59 LPAPSSASEIPKQKOKARLQREVDLYNGMCLQGPAGVGRDGPANGIPETPGIPGR 118  
QY 118 DGRKRGKGLRSEFSESWTPNPKQCSWSLNGIDLGKIAECTPTKMSNSALRYLFGS 177  
DB 119 DGRKRGKGLRSEFSESWTPNPKQCSWSLNGIDLGKIAECTPTKMSNSALRYLFGS 178  
QY 178 SLRLKRNACCQRMWTFNGACSGPLPIEAIYLLDOGSPELNTNHRSTSVEGLCEG 237  
DB 179 SLRLKRNACCQRMWTFNGACSGPLPIEAIYLLDOGSPELNTNHRSTSVEGLCEG 238  
QY 238 IGAGLVDAIWMGTCSDPYKGDASTGMNSVSRILIEELPK 277  
DB 239 IGAGLVDAIWMGTCSDPYKGDASTGMNSVSRILIEELPK 278

RESULT 10

AAM25746 AAM25746 standard; protein; 278 AA.

AC AAM25746;

DT 16-OCT-2001 (first entry)

DE Human protein sequence SEQ ID NO:1261.

XX Human: cancer; HIV infection; human immunodeficiency virus;  
XX antiinflammatory; antirheumatic; antiarthritic; immunosuppressive;  
XX antibacterial; endocrine; cardiant; central nervous system; vitruide;  
XX anti-HIV; fungicide; antimutagen; cardiovascular; antinaemic; anaemia;  
XX antiagregant; haemostatic; vulnerary; anticancer; osteopathic; eczema;  
XX dermatological; antiallergic; antiasbatic; antidiabetic; cytostatic;  
XX neuroprotective; antidepressant; nootropic; antiparkinsonian; infection;  
XX immunostimulant; gene therapy; antisense therapy; vaccine; inflammation;  
XX antinaephylactic; rheumatoid arthritis; septic shock; pancreatitis;  
XX cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmunity;  
XX genetic disease; haematopoietic disorder; platelet disorder; asthma;  
XX thrombocytopaenia; osteoporosis; severe combined immunodeficiency;  
XX allergic rhinitis; diabetes; multiple sclerosis; depression;  
XX Alzheimer's disease; Parkinson's disease; neurodegenerative disorder;  
XX neurological disorder.

XX Homo sapiens.

XX OS

XX

PN W0200153455-A2.

PD 26-JUL-2001.

PF 22-DEC-2000; 2000WO-US035017.

PR 23-DEC-1999; 99US-00471275.

PR 21-JAN-2000; 2000US-00488725.

PR 25-APR-2000; 2000US-00552317.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac R;

DR WPI, 2001-457603/49.

DR N-PSDB; AAH99687.

PT Isolated human polynucleotides encoding polypeptides, useful for the

PS Claim 20, Page 260; 1217p; English.

CC AAH99166 to AAH99904 encode the human proteins given in AAM25225 to  
CC AAM25963. The proteins can have activities based on the tissues and cells  
CC they are expressed in, such as: antiinflammatory; antirheumatic;  
CC antirheumatic; immunosuppressive; antibacterial; endocrine; cardiant;  
CC central nervous system; vitruide; anti-HIV; fungicide; antimutagen;  
CC cardiovascular; antinaemic; antiagregant; haemostatic; vulnerary;  
CC anticancer; osteopathic; dermatological; antiallergic; antiasbatic;  
CC antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic;  
CC antiparkinsonian; and immunostimulant. The proteins and polynucleotides  
CC encoding them can be used in gene therapy, antisense therapy and vaccine  
CC production. The proteins and polynucleotides are useful for screening for  
CC agonists or antagonists of a protein and for the treatment and diagnosis  
CC of disorders associated with the activity of a protein e.g. inflammation,  
CC rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction,  
CC neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal  
CC infections, autoimmunity, genetic diseases, haematopoietic disorders,  
CC anaemia, platelet disorders, thrombocytopaenia, wounds, burns, ulcers,  
CC osteoporosis, severe combined immunodeficiency, eczema, allergic  
CC rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression,  
CC Alzheimer's disease, Parkinson's disease, neurodegenerative and  
CC neurological disorders

SQ Sequence 278 AA:

Query Match 84.0%; Score 1239.5; DB 4; Length 278;  
Best Local Similarity 90.9%; Pred. No. 1.4e-115;  
Matches 239; Conservative 5; Mismatches 14; Indels 5; Gaps 2;

QY 15 GLCRPLCLLCAQSOLPHTMHPOGRAASPOLLGLPLVLLLLQLSAPSSASSENPKYKOKA 74  
DB 21 GLLR---IQLRARAPGAMRPGAPASPORLRG--LLLLLLQLPAPSSASEIPKQKOKA 75  
QY 75 LQREVDLYNGMCLQGPAGVGRDGPANGIPETPGIPGRDGRKRGKGLRSEFSE 134  
DB 76 LQREVDLYNGMCLQGPAGVGRDGPANGIPETPGIPGRDGRKRGKGLRSEFSE 135  
QY 135 SWTPNPKQCSWSLNGYDGLGKIACTPTKMSNSALRYLFGSGSLRLKRNACCQRMWTF 194  
DB 136 SWTPNPKQCSWSLNGYDGLGKIACTPTKMSNSALRYLFGSGSLRLKRNACCQRMWTF 195  
QY 195 FNGACSGPLPIEAIYLLDOGSPELNTNHRSTSVEGLCEGIGAGLVDAIWMGTCS 254  
DB 196 FNGACSGPLPIEAIYLLDOGSPELNTNHRSTSVEGLCEGIGAGLVDAIWMGTCS 255  
QY 255 YPKGDASTGMNSVSRILIEELPK 277  
DB 256 YPKGDASTGMNSVSRILIEELPK 278

RESULT 11

ABP68637



```
ID ABP68637 standard; protein; 278 AA.
XX
XX ABP68637;
XX
XX 14-JAN-2003 (first entry)
XX
XX Human pancreatic cancer expressed protein SEQ ID NO 4560.
XX
XX Human; pancreas; cancer; gene therapy; vaccine; immunostimulant;
XX cytostatic; tumour.
XX
XX Homo sapiens.
XX
XX WO200260317-A2.
XX
XX 08-AUG-2002.
XX
XX 30-JAN-2002; 2002WO-US002781.
XX
XX 30-JAN-2001; 2001US-0265305P.
XX
XX 31-JAN-2001; 2001US-0265682P.
XX
XX 09-FEB-2001; 2001US-0267568P.
XX
XX 21-MAR-2001; 2001US-0278651P.
XX
XX 28-APR-2001; 2001US-0287112P.
XX
XX 16-MAY-2001; 2001US-0291631P.
XX
XX 12-JUL-2001; 2001US-0305484P.
XX
XX 20-AUG-2001; 2001US-0313999P.
XX
XX 27-NOV-2001; 2001US-0333626P.
XX
XX (CORI-) CORIXA CORP.
XX
XX Benson DR, Kalos MD, Lodes MJ, Persing DH, Hepler WT, Jiang Y;
XX
XX WPI, 2002-627435/67.
XX
XX N-PSDB; ABV99145.
XX
XX New isolated polynucleotide and pancreatic tumor polypeptides, useful for
XX diagnosing, preventing and/or treating cancer, particularly pancreatic
XX cancer.
XX
XX Claim 2; SEQ ID NO 4560; 300bp + Sequence Listing; English.
XX
XX The invention relates to an isolated polynucleotide (I) comprising: (a)
XX any of a group of over 4000 nucleotide sequences (ABV94628-ABV99145); (b)
XX complements of (a); (c) sequences consisting of at least 20 contiguous
XX residues of (a); (d) sequences that hybridize to (a), under moderately
XX stringent conditions; (e) sequences having at least 75% or 90% identity
XX to (a); or (f) degenerate variants of (a). Polypeptides (ABP68596-
XX ABP68637) encoded by (I) and oligonucleotide can be used to detect cancer
XX in a patient and compositions comprising polypeptides, polynucleotides,
XX antibodies, fusion proteins, T cell populations and antigen presenting
XX cells expressing the polypeptide are useful in treating pancreatic cancer
XX and stimulating an immune response. The polynucleotides can be used as
XX probes or primers for nucleic acid hybridisation, in the design and
XX preparation of ribozyme molecules for inhibiting expression of the tumour
XX polypeptides and proteins in the tumour cells, in vaccines and for gene
XX therapy. Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at http://wipo.int/pub/published\_pct\_sequences
XX
XX Sequence 278 AA;
XX
XX
XX Query Match 83.6%; Score 1234.5; DB 5; Length 278;
XX Best Local Similarity 86.4%; Pred. No. 4.4e-115;
XX Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;
XX
XX 1 MRPAAL---GOTLSRAGLCRLPCLILCASQLPHTWHPGGRASPOLLGLFVLLILLQ 57
XX :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
XX 1 MGPAAASERGGADADHVPLGLRLQLRAAROGAMRPGGPAASPRRLG--LILLILLQ 58
XX
XX 58 LSAPSSASNPYKQKALIRQREVDLYNGMCLQGAGVPGRDGSGANGIPETPGIPGR 117
XX LPAPSASSEIRPKGKQAKQREVRVDLYNGMCLQGAGVPGRDGSGANGVITPOTPIGR 118
```

```
QY 118 DGFKEKGECLESPFEESWTPNYKQCSWSLNYGIDLGKIAECTTKRNSALFVPSG 177
XX
XX 119 DGFKEKGECLESPFEESWTPNYKQCSWSLNYGIDLGKIAECTTKRNSALFVPSG 178
DB
QY 178 SLRLKCRNACCORWFTFNAGACSGPLPLEAIIYDQSSPELINSTINIHRTSSVGLCEG 237
DB 179 SLRLKCRNACCORWFTFNAGACSGPLPLEAIIYDQSSPELINSTINIHRTSSVGLCEG 238
QY 238 IGAGLVDAIWIWGTCSDPKGDASTGMNSVSRILIEELPK 277
DB 239 IGAGLVDAIWIWGTCSDPKGDASTGMNSVSRILIEELPK 278
XX
XX RESULT 12
XX ABG78939
XX ID ABG78939 standard; protein; 278 AA.
XX
XX AC ABG78939;
XX
XX 15-NOV-2002 (first entry)
XX
XX Human breast tumour polypeptide #30.
XX
XX Human breast tumour protein; breast cancer; cytostatic; vaccine.
XX
XX Homo sapiens.
XX
XX US2002085998-A1.
XX
XX 04-JUL-2002.
XX
XX 13-APR-2001; 2001US-00834759.
XX
XX 28-DEC-1998; 98US-00222575.
XX
XX 02-APR-1999; 99US-00285480.
XX
XX 23-JUN-1999; 99US-00339338.
XX
XX 02-SEP-1999; 99US-00389681.
XX
XX 03-NOV-1999; 99US-00433826.
XX
XX 17-APR-2000; 2000US-00551621.
XX
XX 08-JUN-2000; 2000US-00590751.
XX
XX 22-JUN-2000; 2000US-00604287.
XX
XX 20-JUL-2000; 2000US-00620405.
XX
XX (CORI-) CORIXA CORP.
XX
XX Jiang Y, Dillon DC, Mitcham JL, Xu J, Harlocker SL, Hepler WT;
XX Henderson RA;
XX
XX WPI; 2002-635657/68.
XX
XX N-PSDB; ABS64030, ABS64031.
XX
XX Novel breast cancer polynucleotides and polypeptides encoded by the
XX polynucleotides, useful for detecting the presence of breast cancer in a
XX patient, and in pharmaceutical compositions, for treating breast cancer.
XX
XX Claim 2; Page 236-237; 247bp; English.
XX
XX The invention relates to an isolated breast tumour polynucleotide and the
XX polypeptide it encodes. The polynucleotide and polypeptide are useful for
XX detecting the presence of breast cancer in a patient, and in
XX pharmaceutical compositions for treating breast cancer. The sequences are
XX useful for stimulating an immune response in a patient and can therefore
XX be used in production of vaccines. The sequences are also useful for
XX detecting the presence of a cancer in a patient, by obtaining a
XX biological sample from the patient, contacting the biological sample with
XX a composition of the invention and detecting the amount of polynucleotide
XX that hybridizes to the sample. This sequence represents a human breast
XX tumour polypeptide of the invention
XX
XX Sequence 278 AA;
XX
XX
XX Query Match 83.6%; Score 1234.5; DB 5; Length 278;
```



Best Local Similarity 86.4%; Pred. No. 4.4e-115;  
Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

```
QY 1 MRPAEL---GQTSRAGLCRPCLLCASQLPHTMHPGSAASPOLLLGLFVLLLLQ 57
   1 MOPAAASERGGADADHVPILGLRLQLRAARQGMARPGQPAASPGRLRG--LILLILLQ 58
Db 58 LSAPSSASENPKVKQKALIRQREVDLYNGMCLQGPAGVGRDGPANGIPGTPGIPGR 117
   59 LPAPSSASEIPKQKQKQALRQREVDLYNGMCLQGPAGVGRDGPANAVIPGTPGIPGR 118
QY 118 DGFKEGKEGECLEARSFESWTNPYKQCSWSLNYGIDLGKIAECTFTMRNSALRYLFSG 177
   119 DGFKEGKEGECLEARSFESWTNPYKQCSWSLNYGIDLGKIAECTFTMRNSALRYLFSG 178
Db 178 SLRLKRNACCQKQWYFTFNGAECGSLPIEAIYYLDQSPELNSTINIHRTSSVEGLCEG 237
   179 SLRLKRNACCQKQWYFTFNGAECGSLPIEAIYYLDQSPENNSTINIHRTSSVEGLCEG 238
QY 238 IGAGLVDAIIVGTCSDYPRGDASTGMNSVSRILIIIEELPK 277
   239 IGAGLVDAIIVGTCSDYPRGDASTGMNSVSRILIIIEELPK 278
Db
```

## RESULT 13

ABJ37762  
ID ABJ37762 standard; protein; 278 AA.

AC ABJ37762;

DT 15-MAY-2003 (first entry)

DE Human tumour-related protein - SEQ ID NO 515.

KW Human; vaccine; gene therapy; T cell stimulation; T cell expansion;  
tumour; breast cancer; cancer; immune response stimulation.

OS Homo sapiens.

PN WO200283956-A1.

PD 24-OCT-2002.

PF 15-APR-2002; 2002WO-US012378.

PR 13-APR-2001; 2001US-00834759.

PR 07-DEC-2001; 2001US-00007805.

PR 13-FEB-2002; 2002US-00076622.

XX (CORI-) CORIXA CORP.

PI Houghton RL, Sleath PR, Persing DH, Jiang Y, Dillon DC;  
Mitchem JL, Xu J, Harlocker SL, Hepler WT, Henderson RA, Fanger GR;  
Vedvick TS, McNeill PD, Durham M;

DR WPI; 2003-103376/09.

PT New polypeptide and polynucleotide useful for stimulating and/or  
expanding T cells specific for a tumor protein and treating breast  
cancer.

PS Example 1; Page 322; 375bp; English.

CC The invention comprises a method of stimulating and/or expanding T cells  
specific for a tumor protein. The invention further comprises human  
nucleic acids and proteins that are associated with tumors (e.g. breast  
cancer). The method and sequences of the invention are useful for  
stimulating and/or expanding T cells specific for a tumor protein,  
detecting the presence of cancer, stimulating an immune response in a  
patient and treating breast cancer. The present amino acid sequence  
represents a human tumour-related protein

XX Sequence 278 AA;

Query Match 83.6%; Score 1234.5; DB 6; Length 278;  
Best Local Similarity 86.4%; Pred. No. 4.4e-115;  
Matches 242; Conservative 6; Mismatches 27; Indels 5; Gaps 2;

```
QY 1 MRPAEL---GQTSRAGLCRPCLLCASQLPHTMHPGSAASPOLLLGLFVLLLLQ 57
   1 MOPAAASERGGADADHVPILGLRLQLRAARQGMARPGQPAASPGRLRG--LILLILLQ 58
Db 58 LSAPSSASENPKVKQKALIRQREVDLYNGMCLQGPAGVGRDGPANGIPGTPGIPGR 117
   59 LPAPSSASEIPKQKQKQALRQREVDLYNGMCLQGPAGVGRDGPANAVIPGTPGIPGR 118
QY 118 DGFKEGKEGECLEARSFESWTNPYKQCSWSLNYGIDLGKIAECTFTMRNSALRYLFSG 177
   119 DGFKEGKEGECLEARSFESWTNPYKQCSWSLNYGIDLGKIAECTFTMRNSALRYLFSG 178
Db 178 SLRLKRNACCQKQWYFTFNGAECGSLPIEAIYYLDQSPELNSTINIHRTSSVEGLCEG 237
   179 SLRLKRNACCQKQWYFTFNGAECGSLPIEAIYYLDQSPENNSTINIHRTSSVEGLCEG 238
QY 238 IGAGLVDAIIVGTCSDYPRGDASTGMNSVSRILIIIEELPK 277
   239 IGAGLVDAIIVGTCSDYPRGDASTGMNSVSRILIIIEELPK 278
Db
```

## RESULT 14

ADL93177  
ID ADL93177 standard; protein; 278 AA.

AC ADL93177;

DT 20-MAY-2004 (first entry)

DE Human breast cancer-associated polypeptide #27.

KW gene therapy; protein therapy; vaccine; breast cancer; cancer; human.

OS Homo sapiens.

PN US2003166022-A1.

PD 04-SEP-2003.

PF 15-APR-2002; 2002US-00124805.

PR 28-DEC-1998; 98US-00222575.

PR 02-APR-1999; 99US-00285480.

PR 23-JUN-1999; 99US-0033838.

PR 02-SEP-1999; 99US-00388681.

PR 03-NOV-1999; 99US-00433826.

PR 17-APR-2000; 2000US-0051621.

PR 08-JUN-2000; 2000US-00590751.

PR 22-JUN-2000; 2000US-00604287.

PR 20-JUL-2000; 2000US-00620405.

PR 13-APR-2001; 2001US-00834759.

PR 07-DEC-2001; 2001US-00007805.

PR 13-FEB-2002; 2002US-00076622.

XX (CORI-) CORIXA CORP.

PI Houghton RL, Sleath PR, Persing DH;

DR WPI; 2003-874918/81.

DR N-PSDB; ADL93174.

PT An isolated oncogenic polypeptide useful for preventing, diagnosing and  
treating breast cancer.

PS Example 1; SEQ ID NO 515; 294bp; English.

CC The invention relates to an isolated breast cancer-associated  
polypeptide. The polypeptide may be used for the diagnosis and treatment

CC of breast cancers. The methods are useful for detecting the presence of a  
CC cancer in a patient and treating a cancer in a patient. The present  
CC sequence represents the amino acid sequence of a human breast cancer-  
associated polypeptide.

**SQ** Sequence 278 AA;

Query Match	83.6%;	Score 1234.5;	DB 7;	length 278;
Best Local Similarity	86.4%;	Pred. No. 4.4e-115;		
Matches 242;	Conservative	6;	Mismatches 27;	Indels 5; Gaps 2;

QY	MRPAALD --- GQLTSBAGCRPLCLLLCLCSOLPHMTHEQGRASQULLGLFLVLLILQ	57
Db	1 MOPPAASERGGADADHVPLLLGLLRQLQRAARPGAMRQGPASQRLRG - LILLILLQ	58
QY	LSAPSSASENPKVKOKALIROREVLDLYNGMCLQGPAGVPGDQSPGANGIPGPIGR	117
Db	59 LPAPSSASEIPKXOKAQLRQREVDLYNGMCLQGPAGVPGDQSPGANVPGPIGR	118
QY	DGKGEHGEGLRASEFEESWTPNYKQCSWLSLYGDIKIAECPTKRSALRYLPSG	177
Db	119 DGEFGEGBECLRASEFEESWTPNYKQCSWLSLYGDIKIALCPTKRRSALRYLPSG	178
QY	SLRLKCNACCCQRYWTFENGAECGSLPIEALITYLDQSPBLNSTINIHRTSVYGLCEG	237
Db	179 SLRLKCNACCCQRYWTFENGAECGSLPIEALITYLDQSPBLNSTINIHRTSVYGLCEG	238
QY	IGAGLDVVAIYWGTCSDPYKGDASTGMNSVSIILIEELPK	277
Db	239 IGAGLDVVAIYWGTCSDPYKGDASTGMNSVSIILIEELPK	278

## RESULT 15

ID AAB36667 standard; protein; 243 AA.

AC AAB36667;

DT 14-MAR-2001 (first entry)

DE Human secretory protein TGC-628 SEQ ID NO:7.

KM Human secretory protein; cancer; immune disease; infectious disease;  
KM lung function disorder; liver function disorder; antiinflammacy;  
KM gastrointestinal disorder; cytosatic; haematopoietic; anticoagulant;  
KM immunomodulatory; hepatotropic; cell proliferation-stimulant;  
KM migratory agent; cell differentiation-inducer.

OS Homo sapiens.

PN WO200071581-A1.

PD 30-NOV-2000.

PF 19-MAY-2000; 2000WO-JP003221.

PR 20-MAY-1999; 99JP-00140229.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Itoh Y, Mogi S, Tanaka H, Ohkubo S, Ogi K;

DR WPI; 2001-032023/04.

1. **Introduction**  
 2. **Methodology**  
 3. **Results**  
 4. **Discussion**  
 5. **Conclusion**  
 6. **References**  
 7. **Appendix**  
 8. **Index**  
 9. **Table of Contents**  
 10. **Figure 1**  
 11. **Figure 2**  
 12. **Figure 3**  
 13. **Figure 4**  
 14. **Figure 5**  
 15. **Figure 6**  
 16. **Figure 7**  
 17. **Figure 8**  
 18. **Figure 9**  
 19. **Figure 10**  
 20. **Figure 11**  
 21. **Figure 12**  
 22. **Figure 13**  
 23. **Figure 14**  
 24. **Figure 15**  
 25. **Figure 16**  
 26. **Figure 17**  
 27. **Figure 18**  
 28. **Figure 19**  
 29. **Figure 20**  
 30. **Figure 21**  
 31. **Figure 22**  
 32. **Figure 23**  
 33. **Figure 24**  
 34. **Figure 25**  
 35. **Figure 26**  
 36. **Figure 27**  
 37. **Figure 28**  
 38. **Figure 29**  
 39. **Figure 30**  
 40. **Figure 31**  
 41. **Figure 32**  
 42. **Figure 33**  
 43. **Figure 34**  
 44. **Figure 35**  
 45. **Figure 36**  
 46. **Figure 37**  
 47. **Figure 38**  
 48. **Figure 39**  
 49. **Figure 40**  
 50. **Figure 41**  
 51. **Figure 42**  
 52. **Figure 43**  
 53. **Figure 44**  
 54. **Figure 45**  
 55. **Figure 46**  
 56. **Figure 47**  
 57. **Figure 48**  
 58. **Figure 49**  
 59. **Figure 50**  
 60. **Figure 51**  
 61. **Figure 52**  
 62. **Figure 53**  
 63. **Figure 54**  
 64. **Figure 55**  
 65. **Figure 56**  
 66. **Figure 57**  
 67. **Figure 58**  
 68. **Figure 59**  
 69. **Figure 60**  
 70. **Figure 61**  
 71. **Figure 62**  
 72. **Figure 63**  
 73. **Figure 64**  
 74. **Figure 65**  
 75. **Figure 66**  
 76. **Figure 67**  
 77. **Figure 68**  
 78. **Figure 69**  
 79. **Figure 70**  
 80. **Figure 71**  
 81. **Figure 72**  
 82. **Figure 73**  
 83. **Figure 74**  
 84. **Figure 75**  
 85. **Figure 76**  
 86. **Figure 77**  
 87. **Figure 78**  
 88. **Figure 79**  
 89. **Figure 80**  
 90. **Figure 81**  
 91. **Figure 82**  
 92. **Figure 83**  
 93. **Figure 84**  
 94. **Figure 85**  
 95. **Figure 86**  
 96. **Figure 87**  
 97. **Figure 88**  
 98. **Figure 89**  
 99. **Figure 90**  
 100. **Figure 91**  
 101. **Figure 92**  
 102. **Figure 93**  
 103. **Figure 94**  
 104. **Figure 95**  
 105. **Figure 96**  
 106. **Figure 97**  
 107. **Figure 98**  
 108. **Figure 99**  
 109. **Figure 100**  
 110. **Figure 101**  
 111. **Figure 102**  
 112. **Figure 103**  
 113. **Figure 104**  
 114. **Figure 105**  
 115. **Figure 106**  
 116. **Figure 107**  
 117. **Figure 108**  
 118. **Figure 109**  
 119. **Figure 110**  
 120. **Figure 111**  
 121. **Figure 112**  
 122. **Figure 113**  
 123. **Figure 114**  
 124. **Figure 115**  
 125. **Figure 116**  
 126. **Figure 117**  
 127. **Figure 118**  
 128. **Figure 119**  
 129. **Figure 120**  
 130. **Figure 121**  
 131. **Figure 122**  
 132. **Figure 123**  
 133. **Figure 124**  
 134. **Figure 125**  
 135. **Figure 126**  
 136. **Figure 127**  
 137. **Figure 128**  
 138. **Figure 129**  
 139. **Figure 130**  
 140. **Figure 131**  
 141. **Figure 132**  
 142. **Figure 133**  
 143. **Figure 134**  
 144. **Figure 135**  
 145. **Figure 136**  
 146. **Figure 137**  
 147. **Figure 138**  
 148. **Figure 139**  
 149. **Figure 140**  
 150. **Figure 141**  
 151. **Figure 142**  
 152. **Figure 143**  
 153. **Figure 144**  
 154. **Figure 145**  
 155. **Figure 146**  
 156. **Figure 147**  
 157. **Figure 148**  
 158. **Figure 149**  
 159. **Figure 150**  
 160. **Figure 151**  
 161. **Figure 152**  
 162. **Figure 153**  
 163. **Figure 154**  
 164. **Figure 155**  
 165. **Figure 156**  
 166. **Figure 157**  
 167. **Figure 158**  
 168. **Figure 159**  
 169. **Figure 160**  
 170. **Figure 161**  
 171. **Figure 162**  
 172. **Figure 163**  
 173. **Figure 164**  
 174. **Figure 165**  
 175. **Figure 166**  
 176. **Figure 167**  
 177. **Figure 168**  
 178. **Figure 169**  
 179. **Figure 170**  
 180. **Figure 171**  
 181. **Figure 172**  
 182. **Figure 173**  
 183. **Figure 174**  
 184. **Figure 175**  
 185. **Figure 176**  
 186. **Figure 177**  
 187. **Figure 178**  
 188. **Figure 179**  
 189. **Figure 180**  
 190. **Figure 181**  
 191. **Figure 182**  
 192. **Figure 183**  
 193. **Figure 184**  
 194. **Figure 185**  
 195. **Figure 186**  
 196. **Figure 187**  
 197. **Figure 188**  
 198. **Figure 189**  
 199. **Figure 190**  
 200. **Figure 191**  
 201. **Figure 192**  
 202. **Figure 193**  
 203. **Figure 194**  
 204. **Figure 195**  
 205. **Figure 196**  
 206. **Figure 197**  
 207. **Figure 198**  
 208. **Figure 199**  
 209. **Figure 200**  
 210. **Figure 201**  
 211. **Figure 202**  
 212. **Figure 203**  
 213. **Figure 204**  
 214. **Figure 205**  
 215. **Figure 206**  
 216. **Figure 207**  
 217. **Figure 208**

PT Novel secretory protein and its salt with e.g. anti-cancer, anti-PT inflammatory and hematopoietic, effects, applicable as drugs in remedies PT and preventives to treat diseases like cancer and immune diseases.

PS Claim 1; Page 89-90; 122pp; Japanese.

CC AAC90701 to AAC90715 encode the human secretory proteins given in  
CC AAB36661 to AAB36675. The proteins can have cytosstatic, anti-

CC inflammatory, haematopoietic, anti-coagulant, immunomodulatory and  
CC hepatotropic activities, and can be used as cell migratory agents, cell  
CC proliferation stimulants and cell differentiation-inducers. The proteins  
CC are useful in the treatment and prevention of diseases such as cancer,  
CC lung function disorder, liver function disorder, gastrointestinal  
CC disorder and immune diseases. AAC930716 to AAC930785 represent PCR primers  
CC which are used in the exemplification of the present invention

Sequence 243 AA;

Query Match	83.0%;	Score 1225;	DB 4;	Length 243;
Best Local Similarity	94.7%;	Pred. No. 3.3e-114;		
Matches 232; Conservative	3;	Mismatches 8;	Indels 2;	Gaps 1.

QY	33	MHHQGAASPOLLIIGLFTLLILLQLSAPSSASENPKYKQKALRROREVDLYNGMCLQG	92
Db	1	MRQGGPAAEFQRLRG--LLELLILLQLPAPSSASELPRKQKQALRQREVDLYNGMCLQG	58
QY	93	PAGVPRDGSPGANGIPGTPGI PGDGFKGXGEKCLRESFEESWTPYNYKQCSWSLNYGI	152
Db	59	PAGVPRDGSPPGANGIPGTPGI PGDRDGFKGXGEKCLRESFEESWTPYNYKQCSWSLNYGI	118
QY	153	DLKRIACEPTTKRRSSALRYLFGSLILKCRNACCQWYTFEPNGACSGPLPEALIIYL	212
Db	119	DLGLIACEFTTKRRSSALRYLFGSLILKCRNACCQWYTFEPNGACSGPLPEALIIYL	178
QY	213	DQGSPELNTSTINIHRTSSVEGLCEGIGAGLYDVAIIVWGTCSDYPKGDASTGMSVSRIII	272
Db	179	DQGSPEMNSTINIHRTSSVEGLCEGIGAGLYDVAIIVWGTCSDYPKGDASTGMSVSRIII	238
QY	273	ELLPK 277	
Db	239	ELLPK 243	

Search completed: August 4, 2005, 04:16:46

Job time : 110.817 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:07:24 ; Search time 101.024 Seconds  
(without alignments)  
1404.088 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAALGQTLRSAGLCRPL.....GDASTGMSVSRILIELPK 277

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1307	88.6	245	2	08CG08
2	1292	87.5	245	2	08DND6
3	1225	83.0	243	2	096CG8
4	1216	82.4	243	2	06UW91
5	1063	72.0	232	2	081X63
6	932	63.1	231	2	06AXL0
7	133	9.0	287	2	07Z513
8	133	9.0	287	2	08CFR0
9	132.5	8.9	555	2	08K036
10	131.5	8.9	312	2	08CHX9
11	130.5	8.8	289	2	08BVD7
12	130.5	8.8	551	2	06ZM13
13	129.5	8.8	1669	2	09QZ80
14	127.5	8.6	289	2	018799
15	126	8.5	305	2	09UDP6
16	125.5	8.5	291	2	09NAR3
17	125.5	8.5	298	2	018286
18	125.5	8.5	1752	2	007265
19	125	8.5	717	2	09NOS2
20	125	8.5	1163	2	08N6U4
21	125	8.5	1466	1	CA13 HUMAN
22	124.5	8.4	240	2	06LAJ5
23	124.5	8.4	326	2	0677V9
24	124.5	8.4	358	2	06MEY7
25	124.5	8.4	1336	2	06RZ41
26	124.5	8.4	1516	1	CA1H HUMAN
27	124.5	8.4	1516	2	06RZ39
28	124.5	8.4	1751	2	06RZ40
29	123.5	8.4	311	2	001904
30	123.5	8.4	950	2	086Y22
31	123.5	8.4	957	2	096P44

32	123.5	8.4	957	2	09H0V3	09H0V3	homo sapien
33	123	8.3	1472	2	090ZAO	090ZAO	gallus gall
34	122.5	8.3	282	2	08WFP6	08WFP6	suberites d
35	122.5	8.3	310	2	020282	020282	caenorhabdi
36	122.5	8.3	546	2	06P7U1	06P7U1	mus musculu
37	122.5	8.3	920	2	078BC6	078BC6	mus sp. typ
38	122.5	8.3	1549	2	060444	060444	cricetulus
39	122.5	8.3	2944	2	063870	063870	mus musculu
40	122	8.3	178	2	07M2W9	07M2W9	bos taurus
41	122	8.3	549	2	08EMF8	08EMF8	mus musculu
42	122	8.3	747	1	CA12 BOVIN	P02459	bos taurus
43	122	8.3	1284	2	06P159	06P159	homo sapien
44	121.5	8.2	289	1	COT7 HUMAN	09HX12	homo sapien
45	121	8.2	755	2	000261	000261	homo sapien

## ALIGNMENTS

RESULT 1	ID	Q8CG08	PRELIMINARY;	PRT;	245 AA.
AC	Q8CG08				
DT	01-MAR-2003 (Tremblrel. 23, Created)				
DT	01-MAR-2003 (Tremblrel. 23, Last sequence update)				
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)				
DE	Collagen triple helix repeat-containing protein 1.				
GN	Name=Cthrc1;				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RC	SEQUENCE FROM N.A.				
RP	STRAIN=Sprague-Dawley; TISSUE=8 day balloon-injured carotid artery;				
RL	Lehnert W., Moore D.P., Harmon K.U., Mancini M.L., Lindner V.;				
RT	Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AY136824; AAN15748.1; -				
DR	GO: 0005737; Cytoplasm; IEA.				
DR	GO: 0006817; P:phosphate transport; IEA.				
DR	InterPro: IPR008161; C1g helix.				
DR	InterPro: IPR008160; Collagen.				
DR	Pfam: PF01391; Collagen; 1.				
DR	ProDom: PD000007; C1g_helix; 1.				
KW	Collagen.				
SQ	SEQUENCE 245 AA; 26424 MW; 2296FD6DCDBA21F2 CRC64;				
Query Match	88.6%; Score 1307; DB 2; Length 245;				
Best Local Similarity	100.0%; Pred. No. 1.8e-106;				
Matches	245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
QY	33 MHQGRRAAPOLLGLFLVLLLLQLSAPSSASENKVKOKLIRQEVVDLYNMCLOG 92				
DB	1 MHQGRRAAPOLLGLFLVLLLLQLSAPSSASENKVKOKLIRQEVVDLYNMCLOG 60				
QY	93 PAVGPRGDSRGANGIPGTPGIPGRGFGKGEKCELRFEESWTNYKQCSSSLNYGI 152				
DB	61 PAVGPRGDSRGANGIPGTPGIPGRGFGKGEKCELRFEESWTNYKQCSSSLNYGI 120				
QY	153 DLGKIACEFTKRSNSALRVLPFSGSLRLKCRNACQRYFTFNGAEGSGPLPEAIYLY 212				
DB	121 DLGKIACEFTKRSNSALRVLPFSGSLRLKCRNACQRYFTFNGAEGSGPLPEAIYLY 180				
QY	213 DQSPPELNTINIHRTSVYGLCEGIGAGLVDAIVWGTCSYDYPKGDASTGMSVSRIII 272				
DB	181 DQSPPELNTINIHRTSVYGLCEGIGAGLVDAIVWGTCSYDYPKGDASTGMSVSRIII 240				
QY	273 EELPK 277				
DB	241 EELPK 245				
RESULT 2					

09D1D6  
ID 09D1D6 PRELIMINARY; PRT; 245 AA.  
AC 09D1D6;  
DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched library, clone:1110014B07 product:hypothetical Collagen triple helix repeat containing protein, full insert sequence.  
DE repeat containing protein, full insert sequence.  
GN Name=Ctrnc1;  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
RA Carninci P., Hayashizaki Y.;  
RT "High-efficiency full-length cDNA cloning.";  
RL Meth. Enzymol. 303:19-44(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
RA RIKEN FANTOM Consortium;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RT The FANTOM Consortium.  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;  
RT "Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.";  
RL Genome Res. 10:1617-1630(2000).  
RN [5]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P., Kono H., Akiyama J., Nishi K., Kiteunai T., Tashiro H., Itoh M., Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A., Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K., Fujiwaki S., Inoue K., Togawa K., Izawa M., Ohara E., Watanabe M., Yonekura Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J., Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
RT "RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer.";  
RL Genome Res. 10:1757-1771(2000).  
RN [6]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Whole body;  
RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H., Arikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M., Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoaka T., Hori F., Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H., Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C., Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M., Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K., Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T., Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T., Tejima Y., Toyota T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK003674; BAB22930.1; -.  
DR WGI; WGI:1915638; Ctrnc1.  
DR GO; GO:0005737; Cytoplasm; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008161; C1g helix.  
DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen; 1.  
DR ProDom; PD000007; C1g\_helix; 1.  
KW Collagen, Hypothetical protein.  
SQ SEQUENCE 245 AA; 26460 MW; 14951B87D81B1A0E CRC64;  
Query Match 87.5%; Score 1292; DB 2; Length 245;  
Best Local Similarity 98.8%; Pred. No.3,8e-105;  
Matches 242; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 33 MHPOGRAAPOLLGLFLVLLLLQLSPSSASNPYKQKALIRQREVVDLYNMCILQG 92  
DB 1 MHPOGRAAPOLLGLFLVLLLLQLSPSSASNPYKQKALIRQREVVDLYNMCILQG 60  
QY 93 PAVGVRGRDGSFRANCIPTPTGIPGRDGRKGEKGECLRSFESWTPNPKQCSWSSLYNGI 152  
DB 61 PAVGVRGRDGSFRANCIPTPTGIPGRDGRKGEKGECLRSFESWTPNPKQCSWSSLYNGI 120  
QY 153 DLGKIAECTFTMGKNSALRYLFSGLRLKCRNACCQWYFTPNGAECGSLPIPIATLYL 212  
DB 121 DLGKIAECTFTMGKNSALRYLFSGLRLKCRNACCQWYFTPNGAECGSLPIPIATLYL 180  
QY 213 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAVIAWGTCSDDYPRKDASTGMVNSYRIII 272  
DB 181 DQGSPELNSTINIHRTSSVEGLCEGIGAGLVDAVIAWGTCSDDYPRKDASTGMVNSYRIII 240  
QY 273 EELPK 277  
DB 241 EELPK 245  
RESULT 3  
Q96CG8 PRELIMINARY; PRT; 243 AA.  
ID Q96CG8  
AC Q96CG8;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
DE Collagen triple helix repeat containing 1 (Collagen triple helix repeat-containing protein 1).  
GN Name=CTRNC1;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RX MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold R.A., Grouse L.H., Berger J.G., Klauener R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheffer C.F., Bhat N.K., Hopkins L.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Urdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J., Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Mizny D.M., Sodergren B.J., Lu X., Gibbs R.A., Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywicki M.I., Skalski U., Smalios D.E., Scherch N., Schein J.E., Jones S.J., Maier M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human

RT		and mouse cDNA sequences."
RL		Proc. Natl. Acad. Sci. U.S.A.
RN	(2)	99:16899-16903(2002).
RN	[2]	
RP		SEQUENCE FROM N.A.
RC		TISSUE=Kidney;
RA		Strausberg R.;
RL		Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
RN	[3]	
RP		SEQUENCE FROM N.A.
RC		TISSUE=Aorta smooth muscle;
RA		Lehnert W., Moore D.P., Harmon K.J., Mancini M.L., Lindner V.;
RL		Submitted (JUL-2002) to the EMBL/Genbank/DBJ databasees.
DR	EMBL:	B0142424; AAH14245.1; -
DR	EMBL:	AY136825; AAN15749.1; -
DR	GeneW;	HGNC:18831; CTRC1.
DR	GO;	GO:0005737; Cytoplasm; IEA.
DR	GO;	GO:006817; P-phosphate transport; IEA.
DR	InterPro;	IPIR08161; Clg_helix.
DR	InterPro;	IPIR08160; Collagen.
DR	Pfam;	PF01391; Collagen_1.
DR	Prodrom;	PD000007; Clg_helix; 1.
KV	C collagen.	
SO	SEQUENCE	243 AA; 26224 MW; AllFEBIC66867f9 CRC64;
Query Match	Best Local Similarity	83.0%; Score 1225; DB 2; Length 243;
Matches 232;	Conservative	3; Mis.matches 8; Indels 2; Gaps 1
Oy	33	MHPQARAAPOLLGLPYLLTLLOLSAFSSASSENPKVKOKLIIRREVDLYNGMCLQG 92
Dd	1	MPPQPASPSPORLR--LTLLLTLQLPAPSASELPKQKOCLRERVVLDYNMCLOG 58
Oy	93	PAGVGGRDGPANGIPGTPTGIPIGRDFXGEKGCECIARESFESBWTNTNYQCWSISINYGI 152
Dd	59	PAGVGGRSGSPANGIPLTPTGTIGRNRFGEKEGCARCFEBSFTNTNYQCWSISINYG 118
Oy	153	DLGKTACTCTFYMRNSALRVLFPSGSRLIKRNAACCRNVFFTFNAGECGPLEIAIIYL 212
Dd	119	DLGKAETCTFTMRNSALRVLFPSGSLRIKCENNAACQRNVFTFNAGECGPLEIAIIYL 178
Oy	213	DQSSELNSTNIHTTSVEGLECEGIGAIVAVIWVCSDVPKPDAATGMVSARIIT 272
Dd	179	DQSESNMTSNIHRSTSVEGLECEGIGAELDVVAIVWGCSDPDKPADSTMGVSRIT 238
Oy	273	EELPR 277
Dd	239	EELPR 243
<hr/>		
RESULT 4		
O6UM91 PRELMINARY; FRT; 243 AA.		
ID O6UM91		
AC O6UM91.		
DT 05-JUL-2004 (TREMBLrel. 27, Created)		
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)		
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)		
DE CTNRCl.		
Gn ORNames=UNOT62;		
Os Homo sapiens (Human).		
Oc Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Oc Mammalia; Euteria; Primates; Catarrhini; Homindae; Homo.		
Ox NCBI_TaxId=9606;		
RN [1]		
RP SEQUENCE FROM N.A.		
RX MEDLINE=22887296; PubMed=12975309; DOI=10.1101/gr.1293003;		
RA Clark H.F., Gurley A.L., Abaya E., Baker K., Baldwin D., Brush J.,		
RA Chen Y., Chow B., Choi C., Crowley C., Currell B., Deuel B., Dowd P.,		
RA Eaton D., Foster J., Grimaldi C., Gu Q., Haas P.E., Heidens S.,		
RA Huang L., Kim H.-S., Klimowski L., Jin Y., Johnson S., Lee J.,		
RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,		
RA Seehagiri S., Simone L., Singh U., Smith V., Stinson J., Vargas A.,		
RA Vandlen R., Waterabe C., Wieand D., Woods K., Xie M.H., Yanaura D.,		
RA Yi S., Yu G., Yuan J., Zhang W., Zhang Z., Goddard A., Wood W.I.,		

R	A
RT	"The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment."
RL	Genome Res. 13:2265-2270(2003).
DR	EMBL; AY358914; AAC092273.1; -.
DR	GO; GO:0005737; C:cytosolem; IEA.
DR	InterPro; IPR008161; Clg_helix.
DR	Prodrom; PD000007; Clg_helix; 1.
KW	Collagen.
SQ	SEQUENCE 243 AA; 26266 MW; BCB49AFDBC303BC CRC64;
OY	Query Match 82.4%; Score 1216; DB 2; Length 243; Best Local Similarity 94.3%; Pred. NO. 1.7e-98;
D	Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1
OY	33 MHPOGRAAPOLLGLFLVLLILQLSPPSSASSENPKVKOKLIRORREVVLDLYNGMCLOG  MRPGGPASPRLK--DLLHLIOLPAPSASELPKGKKOQLRORREVDLYNKCLOG 58
D	93 PAVVPGRDGSGANGIPCTPGIPGRDXGKEGECECLARSFESWTPTNYKCSWSLTAYGI 1522 59 PAVVPGRDGSGANVTPECTPGIPGRDXGKEGECECLARSFESWTPTNYKCSWSLTAYGI 118
OY	153 DLGXIAETFTFKMRSNALRYLFSSSLTLCKRNACCOWFFPNAGECGGPIEAIITVL 212 119 DLGXIAECTFTFKMRSNSALRYLFSSLSLTKCRMACCOWFFPFNGACCGPLIEAIITVL 178
OY	213 DGGSPELNSTINIHRTSVSEVEIGDEIGLVDAIVDWGCSCPYPGADASTGNMSVRRIII 272 179 DGGSPEMNSTINHRTSSVBELCBEGICGVADVAILWGTCSDYPKGDASTGNMSVRRIII 238
D	273 EELPK 277     239 EELPK 243
RESULT 5	
ID	Q8IX63 PRELIMINARY; PRT; 232 AA. Q8IX63
AC	Q8IX63
DT	01-MAR-2003 (TrEMBLrel. 23 Created)
DT	01-MAR-2003 (TrEMBLrel. 23 Last sequence update)
DT	01-OCT-2003 (TrEMBLrel. 25, last annotation update)
DE	NTRYCL.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
CX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Sanuki N., Fujiki K., Kanai A., Tanaka Y., Iwata T.; Submitted (JUN-2001) to the EMBL/Genbank/DBJ databases.
DR	EMBL; AF395488; AAC17919.1; -
DR	GO; GO:0005737; C:cytosoleum; IEA.
DR	GO; GO:0006817; P:phosphate transport; IEA.
DR	InterPro; IPR008161; Clg_helix.
DR	InterPro; IPR008160; Collagen.
Pfam:	PF01391; Collagen; 1.
DR	Prodrom; PD000007; clg_helix; 1.
KW	COLLAGEN .
SQ	SEQUENCE 232 AA; 25163 MW; ESPDABC3030483TED CRC64;
OY	Query Match 72.0%; Score 1063; DB 2; Length 232; Best Local Similarity 87.2%; Pred. No. 4.4e-85;
D	Matches 197; Conservative 7; Mismatches 12; Indels 10; Gaps 1
OY	61 PSSASSENPKVKOKLIRORE-----VVDLYNMGCILOGPAVGPDGDSFGANGTTPG 110  :::  :::  :::  ::    ::  3 PPGRISIVKAKRKVKVSRLTNNGPSAFQGILLCGKTNGMICOGPAVGPDGDSFGANGTTPG 62



RP SEQUENCE FROM N.A.  
 RC TISSUE=Eye;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haile F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywiński M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Eye;  
 RC Strausberg R.;  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC040774; AAH40774.1; -  
 DR HSSP; Q60994; 1C28.  
 DR MGD; MGI:3032521; BC040774.  
 DR GO; GO:0005737; C:cytoplasm; IEA.  
 DR GO; GO:0006817; P:phosphate transport; IEA.  
 DR InterPro; IPR001073; Clq.  
 DR InterPro; IPR008160; Collagen.  
 DR InterPro; IPR008983; TNF\_like.  
 DR Pfam; PF00386; Clq; 1.  
 DR DR Pfam; PF01391; Collagen; 1.  
 DR PRINTS; PRO0007; COMPENNTCLO.  
 DR SMART; SM00110; Clq; 1.  
 DR PROSITE; PS01113; Clq; 1.  
 DR Collagen.  
 SQ SEQUENCE 287 AA; 29292 MW; 8FF69EC1C7420415 CRC64;  
 Query Match 9.0%; Score 133; DB 2; Length 287;  
 Best Local Similarity 27.5%; Pred. No. 0.0022;  
 Matches 46; Conservative 16; Mismatches 61; Indels 44; Gaps 3;  
 QY 44 LILGLFVLLILLQLSPSSASENPVKOKALIFOREVVDLYN----- 86  
 Db 1 MALGLIIVPLLIQAAAPGAA-----HTEMLGTCTWICDPYVAAPGAPGAKAPPPGP 54  
 QY 87 -----GMCLOGPAGVPGRDSPGANGIPGTGIPGRDGFKEGEGCLNE 130  
 Db 55 STAALEVMODLSANPPPPFQGPDPGRGKGPGRGPPGPPEPGRGPPGKSGNSGR 114  
 QY 131 SPEESTPNVKQCSMSLANTGIDLGKIAECTFTKASNSMLRLVBSG 177  
 Db 115 GL-----PGLQLTTSAAAGVGVSGGTGGGGDTEGEVTSALSAFSG 156  
 RESULT 9  
 ORK036 PRELIMINARY; PRT; 565 AA.  
 AC 08K036;  
 DT 01-OCT-2002 (TREMBLrel. 22, Created)  
 DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Clqnf7 protein.  
 GN Name=Clqnf7;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=mlx FVB/N; TISSUE=Mammary tumor;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haile F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywiński M.I., Skalska U., Smalins D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=mlx FVB/N; TISSUE=Mammary tumor;  
 RC Strausberg R.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC034164; AAH34164.1; -  
 DR MGD; MGI:1272201; Coll13a1.  
 DR GO; GO:005911; C:intercellular junction; IDA.  
 DR InterPro; IPR008161; Clq\_helix.  
 DR InterPro; IPR008160; Collagen.  
 DR Pfam; PF01391; Collagen; 5.  
 DR Prodom; PD000007; Clq\_helix; 1.  
 DR Collagen.  
 SQ SEQUENCE 565 AA; 56726 MW; DBD3FF99D670195F CRC64;  
 Query Match 9.0%; Score 132.5; DB 2; Length 565;  
 Best Local Similarity 32.7%; Pred. No. 0.0052;  
 Matches 36; Conservative 10; Mismatches 35; Indels 29; Gaps 5;  
 QY 59 SAPSSASENPVKOKALIFOREVVDLYN-----MCLQGPAGVPGRDSPGANG 107  
 Db 402 SGPFGAKGPP-----GKGEHVD-YNCSINEALQEIETLALMGPPGIPGOTGPPGPPG 452  
 QY 108 IP-----CTPGIPGRDGFKEGEGCLRSPFEESTPNVKQCSMSLANTG 151  
 Db 453 TPGQREITGLPPGPHDGDGPPG--KPGDAGWSSRTPRGKWSRNEG 499  
 RESULT 10  
 ORK039 PRELIMINARY; PRT; 312 AA.  
 AC 08K039;  
 DT 01-MAR-2003 (TREMBLrel. 23, Created)  
 DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Clqnf7 protein (Fragment).  
 GN Name=Clqnf7;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,



```
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Ronald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullen S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,
RA Richardson S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RC SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Mammary gland;
RA Straubeberg R.;
RL Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC038308; AAH38308.1; -.
DR HSSP; Q60994; 1C28.
DR MGD; MGI:1925911; C1qtnf7.
DR GO; GO:0005737; C-cytoplasm; IEA.
DR GO; GO:0006817; P-phosphate transport; IEA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR PROSITE; PS01113; C1Q; 1.
DR Collagen.
KW NON TER
FT
SQ SEQUENCE 312 AA; 33162 MW; 1E240A83A5014F43 CRC64;

Query Match 8.9%; Score 131.5; DB 2; Length 312;
Best Local Similarity 32.4%; Pred. No. 0.0033;
Matches 34; Conservative 7; Mismatches 25; Indels 39; Gaps 2;

OY 22 LLLCASQLPHTMHPGSAASPOLLLGLFLVLLLLQLSPSSASBPVKYKQKALINQRAV 81
DB 33 LAICASGQPRANDAKESYSPIRI-----CSIPG----- 61
OY 82 VDLVNGMCLQGPAGVGRDGSFGANGIPGTPGIPGRDGFKEGKE 126
DB 62 -----LPGRPPGPGANGSPGPHRIGLPGRDGRDKRKEGKE 98

RESULT 11
Q8BVD7 PRELIMINARY; PRT; 289 AA.
ID O8BVD7;
AC O8BVD7;
DT 01-MAR-2003 (TEMBLrel. 23, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last annotation update)
DE Mus musculus 16 days embryo lung cDNA, RIKEN full-length enriched
DE library, clone:8430425G24 product:COMPLEMENT-C1Q TUMOR NECROSIS
DE FACTOR-RELATED PROTEIN homolog.
GN Name=C1qtnf7;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6679(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
```

```
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RC SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasaiwagi K.,
RA Fujiwara S., Inoue K., Ozawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Adachi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Aikawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hangaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imocani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamuro T., Yamataka I., Yasunishi A.,
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK078818; BAC37409.1; -.
DR HSSP; Q60994; 1C28.
DR MGD; MGI:1925911; C1qtnf7.
DR GO; GO:0005737; C-cytoplasm; IEA.
DR GO; GO:0006817; P-phosphate transport; IEA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
KW Collagen.
SQ SEQUENCE 289 AA; 30483 MW; 514AC7P2C318832A CRC64;

Query Match 8.8%; Score 130.5; DB 2; Length 289;
Best Local Similarity 32.4%; Pred. No. 0.0037;
Matches 34; Conservative 7; Mismatches 25; Indels 39; Gaps 2;
```



```

QY 22 LILCASOLPHTMHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREV 81
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 10 LALCASQPPNPQAKGSSYSPRYI-----CSIFG----- 38

QY 82 VDLVNGMCLQGPAGVPGRDSPGANGIPGTGPIGPDGPFKGEKGE 126
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 39 -----LPGPPGPGANGSPGHGIGLPGDGRGRGKEKGE 75

RESULT 12
Q6ZMT3 PRELIMINARY; PRT; 551 AA.
AC 06ZMT3;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE Hypothetical protein FLJ23917.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Kawabata A., Hkij T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,
RA Tanaka T., Nakamura Y., Isegai T., Sugano S.;
RA Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AKI72756; BAD18742.1; -.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR003112; Olfac_like.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF02191; OLF; 1.
DR Prodom; PD000007; Clg_helix; 1.
DR SMART; SM00284; OLF; 1.
KW COLLAGEN.
SQ SEQUENCE 551 AA; 58957 MW; CE14A36120DEC818 CRC64;

Query Match 8.8%; Score 130.5; DB 2; Length 551;
Best Local Similarity 30.7%; Pred. No. 0.0076;
Matches 35; Conservative 18; Mismatches 20; Indels 41; Gaps 5;

QY 35 PGRAPSPOLLGLFLVLLLLQLSAPSSASEN-----PKYKOK-----AL 75
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 76 PGRASAPQ-----DPASARKRSHSGEPAPHIRAESHDMMLMTYSM 119

QY 76 IRQREVVDLYN--GMCLQGPAGVPGRDSPGANGIPGTGPIGPDGPFKGEKGE 126
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 120 VPIRVAVDLCNSYKIGLITGSPSP--GPPAGAGLPGHGLDQGPQGPQKGE 170

RESULT 13
Q9QZSO PRELIMINARY; PRT; 1669 AA.
AC 09QZSO;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-MAY-2004 (TReMBLrel. 26, Last annotation update)
DE Alpha 3 collagen IV.
GN Name=Col4a3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA MEDLINE=2000594; PubMed=10534397; DOI=10.1006/geno.1999.5943;
RA Lu W., Phillips C.L., Kilen P.D., Hlaing T., Harrison W.R.,
RA Elder F.F.B., Miner J.H., Overbeek P.A., Meisler M.H.;

```

```

RT "Insertional mutation of the collagen genes Col4a3 and Col4a4 in a
RT mouse model of Alport syndrome.";
RL Genomice 61.113-124(1999).
DR EMBL; AF169387; AAD50449.1; -.
DR PIR; I48302; I48302.
DR HSSP; P02462; 1L1I.
DR MGD; MGI:104688; Col4a3.
DR GO; GO:0005877; C:collagen type IV; ISS.
DR GO; GO:0005615; C:extracellular space; TAS.
DR GO; GO:0005178; F:integrin binding; ISS.
DR GO; GO:0008191; F:metalloendopeptidase inhibitor activity; ISS.
DR GO; GO:0006919; P:caspace activation; ISS.
DR GO; GO:0008283; P:cell proliferation; ISS.
DR GO; GO:0007166; P:cell surface receptor linked signal transdu. .; ISS.
DR GO; GO:0006917; P:induction of apoptosis; ISS.
DR GO; GO:0016525; P:negative regulation of angiogenesis; ISS.
DR InterPro; IPR008161; Clg_helix.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR004829; Csurface antigen.
DR InterPro; IPR001442; Procollagen4_C.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF01413; C4; 2.
DR Pfam; PF01391; Collagen; 21.
DR Prodom; PD000007; Clg_helix; 9.
DR Prodom; PD153432; Csurface antigen; 2.
DR Prodom; PD003923; Procollagen4_C; 2.
DR SMART; SM00111; C4; 2.
DR PROSITE; PS00030; RRM_RNP_1; UNKNOWN_1.
KW COLLAGEN.
SQ SEQUENCE 1669 AA; 161769 MW; 30976E59739A47B2 CRC64;

Query Match 8.8%; Score 129.5; DB 2; Length 1669;
Best Local Similarity 37.0%; Pred. No. 0.033;
Matches 37; Conservative 9; Mismatches 29; Indels 25; Gaps 4;

QY 33 MHPOGRAAPOLLGLFLVLLLLQLSAPSSASENPKVKOKALIRQREVVDLYNGMCL-- 90
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 1 MHSK---TAPRL--VFLLITLLILAAAPVASKGVCVKGK-----GGCLCA 42

QY 91 ----QGPAVPGRDSPGANGIPGTGPIGPDGPFKGEKGE 125
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 43 GTKGEKGEKGVPGSPGPFQKGFPPGEGLPQGPGKSGSG 82

RESULT 14
Q18799 PRELIMINARY; PRT; 289 AA.
AC 018799;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-JUN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Hypothetical protein C53B4.5.
GN ORFName=C53B4.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Berks M.;
RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68215; CAA92453.1; -.
DR PIR; T20177; T20177.
DR WormBase; WBGene00000693; C53B4.5.

```

DR Wormpmed; C53B.5; CE03091.  
DR GO; GO:0055737; C:cytooplasm; IEA.  
DR GO; GO:0042302; E:structural constituent of cuticle; IEA.  
DR GO; GO:0006817; P:phosphate transport; IEA.  
DR InterPro; IPR008160; Collagen.  
DR InterPro; IPR02486; Col\_cuticle\_N.  
DR Pfam; PF01391; Collagen\_2.  
DR Pfam; PF01484; cuticle\_N; 1.  
KM Collagen; Hypothetical protein.  
SQ SEQUENCE 289 AA: 28993 MW: 29996A PFC0044 CRC6;

Query Match	8.6%	Score 127.5;	DB 2;	Length 289;
Best Local Similarity	48.1%	Pred. No. 0.0068;		
Matches 25; Conservative	4;	Mismatches 8;	Indels 15;	Gaps 1;

```
Qy      91 QGPAGVGRDGS-----PGANGIPCTPGIPGRDGEKGEKGC 127
          :| | | |:| | | | | | | | | | | | | | | |
Db     220 RGPAGCGCKDGAQGSGPEKGANGLPGQPGRRGQPGRPQGRDGHPEKGV 271
```

## RESULT 15

ID	Q9UDP6	PRELIMINARY;	PRT;	305 AA.
AC	Q9UDP6			
DT	01-MAY-2000	(TrEMBLrel. 13, Created)		
DT	01-MAY-2000	(TrEMBLrel. 13, Last sequence update)		
DT	01-MAR-2004	(TrEMBLrel. 26, Last annotation update)		
DE	Type XIV collagen (Fragment).			
OS	Homo sapiens (human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
NN	[1]			

RP SEQUENCE.  
RX MEDLINE=95128518; PubMed=7827751; DOI=10.1016/0945-053X(94)90194-5;  
RA Brown J.C., Goldik R., Mann K., Timpl R.;  
EL Matrix Biol. 14:287-295 (1994).

DR PIR; S37749; S37749.  
DR PIR; S46657; S46657.  
DR GO; GO:0005737; C:cytoplasm; IEA.  
DR GO; GO:0006117; P:phosphate transport; IEA.  
DR InterPro; IPR008161; Clg\_helix.  
DR InterPro; IPR008160; Collagen.  
DR Pfam; PF01391; Collagen; 4.  
DR ProDom; PD000007; Clg\_helix; 2.  
DR Collagen.  
SQ SEQUENCE 305 AA; 28934 MW; 2F8CF9423BE111

Query Match	8.5%;	Score 126;	DB 2;	Length 305;
Best Local Similarity	31.9%;	Pred. No. 0.0098;		
Matches 37;	Conservative 11;	Mismatches 34;	Indels 34;	Gaps 5

```
Qy      32 TMHPOGRAASQQLLGLFVLVLLLLQLSAPSSASENPVKYKQALIR--QREYVD----- 83
          :| | | :         ||::||::||:
Db     127 SMSGPGCALGPPGV-----PGAKGERGERGLQSQMVRXARQYAEQLISH 173
```

```

Oy      84  --LYNGM-----CLOGPAGVPERDGSPCA---NGI PCGTPPI PGHDGFKGEKGE 126
          || :      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      174  MARYNQIPSHSSIRTVGDPGCEGRPGSPCAPGEGGPPGCTGPGFPNAGVPGTIPGE 229

```

Search completed: August 4, 2005, 04:21:24  
Job time : 102.024 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 4, 2005, 04:08:16 ; Search time 22.0876 Seconds  
(without alignments)  
1206.653 Million cell updates/sec

Title: US-10-634-108-5

Perfect score: 1476

Sequence: 1 MRPAALGQTLSPAGLCRPL.....GDASTGMSVSRRIIEELK 277

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs; 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	127.5	8.6	289	T20177	hypothetical prote
2	125.5	8.5	298	T27644	hypothetical prote
3	125.5	8.5	1752	A45407	collagen alpha 3(I
4	125	8.5	1466	1 CGHUTL	collagen alpha 1(I
5	123.5	8.4	311	2 T15268	hypothetical prote
6	122.5	8.3	310	2 T29731	hypothetical prote
7	122.5	8.3	920	2 A45748	collagen alpha 1(V
8	122.5	8.3	1549	2 148103	type VII collagen
9	122	8.3	178	2 A39762	collagen alpha 1(X
10	122	8.3	330	2 S46557	collagen alpha 1(X
11	121	8.2	1763	2 S19366	collagen alpha 2(I
12	119.5	8.1	246	2 S29328	complement subcomp
13	119.5	8.1	323	2 A61396	collagen alpha 1(I
14	119.5	8.1	1670	1 CGHUB3	collagen alpha 3(I
15	119.5	8.1	2944	2 A54649	collagen alpha 1(V
16	119	8.1	245	1 C1HUC	complement subcomp
17	119	8.1	298	1 JCI1448	collagen col-34 -
18	119	8.1	299	2 T29956	hypothetical prote
19	119	8.1	1049	1 CGB078	collagen alpha 1(I
20	118	8.0	458	2 T31631	hypothetical prote
21	118	8.0	1669	1 CGMS4B	collagen alpha 1(I
22	118	8.0	1744	2 S40991	collagen alpha 1(I
23	117.5	8.0	324	2 A31920	collagen sct-1 pre
24	117.5	8.0	673	1 CGB06C	collagen alpha 1(I
25	117	7.9	253	2 S49158	complement protein
26	116.5	7.9	58	2 A45526	collagen homolog A
27	116.5	7.9	300	2 T24482	hypothetical prote
28	116	7.9	289	2 T26812	hypothetical prote
29	116	7.9	302	2 T15936	hypothetical prote

30	116	7.9	325	2 T18594	hypothetical prote
31	116	7.9	744	2 S15435	collagen alpha 1(V
32	115.5	7.8	349	2 S08278	macrophage scaveng
33	115.5	7.8	428	2 T24769	hypothetical prote
34	115.5	7.8	453	2 S08276	macrophage scaveng
35	115	7.8	296	2 A31219	collagen 1 - Caeno
36	115	7.8	301	2 B31219	collagen 2 - Caeno
37	115	7.8	744	1 A34246	collagen alpha 1(V
38	115	7.8	744	1 S23298	collagen alpha 1(V
39	115	7.8	1042	1 CGCH15	collagen alpha 1(I
40	115	7.8	1142	2 JX0369	collagen alpha 1(X
41	114.5	7.8	248	1 LNHUP1	pulmonary surfacta
42	114.5	7.8	1496	1 CGHUTV	collagen alpha 2(V
43	114.5	7.8	1497	2 T49607	procollagen type V
44	114	7.7	327	2 T29031	hypothetical prote
45	114	7.7	451	2 A38415	macrophage scaveng

#### ALIGNMENTS

##### RESULT 1

T20177

hypothetical protein C53B4.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T20177

R:Berke, M.

Submitted to the EMBL Data Library, December 1995

A:Reference number: Z19233

A:Accession: T20177

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-289 <MIL>

A:Cross-references: UNIPROT:Q18799; EMBL:Z68215; PIDN:CAA92453.1; GSPDB:GN00022; CESP:ZK

A:Experimental source: clone C53B4

C:Genetics:

A:Gene: CESP:C53B4.5

A:Map position: 4

Query Match

Best Local Similarity 8.6%; Score 127.5; DB 2; Length 289;

Matches 25; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

QY

91 QGPAGVPERDGS-----PGANGIPGTGIPGRDGFGEKGECC 127

DB 220 RGPAGQPKDGAQGGPGEKANGEPQGPGRDGPGRPGGRDGHGEKGVCC 271

##### RESULT 2

T27644

hypothetical protein ZK1010.7 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C:Accession: T27644

R:Gardner, A.

Submitted to the EMBL Data Library, November 1996

A:Reference number: Z20398

A:Accession: T27644

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-228 <MIL>

A:Cross-references: UNIPROT:O18286; EMBL:Z82083; PIDN:CAB04973.1; GSPDB:GN00021; CESP:ZK

A:Experimental source: clone ZK1010

C:Genetics:

A:Gene: CESP:ZK1010.7

A:Map position: 3

A:Introns: 54/3; 236/1

Query Match

Best Local Similarity 8.5%; Score 125.5; DB 2; Length 298;

Matches 30; Conservative 6; Mismatches 29; Indels 11; Gaps 2;



A:Cross-references: GB:5622925; NID:9386425; PIDN:AAJ1397.1; PID:94261637  
R.Chiodo, A.A.; Silience, D.O.; Cole, W.G.; Bateman, J.F.  
Biochem. J. 311, 939-943, 1995  
A:Title: Abnormal type III collagen produced by an exon-17-skipping mutation of the COL3A1 gene  
A:Reference number: 559511; MUID:96067614; PMID:7487954  
A:Accession: S59511  
A:Molecule type: mRNA  
A:Residues: 302-423 <CHT>  
A:Cross-references: GB:579877; NID:91195576; PIDN:AA35615.1; PID:91195577  
R.Seyer, J.M.; Kang, A.H.  
Biochemistry 17, 3404-3411, 1978  
A:Title: Covalent structure of collagen: amino acid sequence of five consecutive CNBr peptides  
A:Reference number: A90414; MUID:79000343; PMID:687591  
A:Accession: A90414  
A:Molecule type: protein  
A:Residues: 399-675, 'N', 677-727 <SEY3>  
A:Experimental source: liver  
R.Lee, B.; Vitale, E.; Superti-Furga, A.; Steinmann, B.; Ramirez, F.  
J. Biol. Chem. 266, 5258-5259, 1991  
A:Title: G to T transversion at position +5 of a splice donor site causes skipping of the first exon in the COL1A1 gene  
A:Reference number: 155349; MUID:91161621; PMID:1672129  
A:Accession: 155349  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 537-605 <LEB>  
A:Cross-references: GB:M59312; NID:9180815; PIDN:AA52041.1; PID:9180816  
R.Seyer, J.M.; Mainardi, C.; Kang, A.H.  
Biochemistry 19, 1583-1589, 1980  
A:Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB5 from type I collagen  
A:Reference number: A90438; MUID:80198282; PMID:6246925  
A:Accession: A90438  
A:Molecule type: protein  
A:Residues: 728-895, 'A', 897-964 <SEY4>  
A:Experimental source: liver  
R.Cole, W.G.; Chiodo, A.A.; Lamande, S.R.; Janeczko, R.; Ramirez, F.; Dahl, H.H.M.; Charney, J.  
J. Biol. Chem. 265, 17070-17077, 1990  
A:Title: A base substitution at a splice site in the COL3A1 gene causes exon skipping and a severe form of Ehlers-Danlos syndrome  
A:Reference number: A38303; MUID:9109133; PMID:2145268  
A:Accession: A38303  
A:Molecule type: mRNA  
A:Residues: 861-1015 <COL>  
A:Cross-references: GB:M05517; GB:M55603; GB:M55227; NID:9180878; PIDN:AA55383.1; PID:9180878  
A:Note: a mutant sequence with 942-977 spliced out from a patient with Ehlers-Danlos syndrome  
R.Mannoo, B.S.; Dalgleish, R.  
Nucleic Acids Res. 16, 2337, 1988  
A:Title: Human pro alpha1(III) collagen: cDNA sequence for the 3' end.  
A:Reference number: S02119; MUID:88189827; PMID:3357782  
A:Accession: S02119  
A:Status: translation not shown  
A:Status: translation not shown  
A:Molecule type: mRNA  
A:Residues: 950-1018, 'Y', 1020-1183, 'S', 1185-1466 <MAN>  
A:Cross-references: EMBL:X06700; NID:930053; PIDN:CAA2886.1; PID:930054  
R.Seyer, J.M.; Kang, A.H.  
Biochemistry 20, 2621-2627, 1981  
A:Title: Covalent structure of collagen: amino acid sequence of alpha1 (III)-CB9 from type I collagen  
A:Reference number: A90446; MUID:81208139; PMID:7016180  
A:Accession: A90446  
A:Molecule type: protein  
A:Residues: 965-979, 'A', 981-984, 'PS', 987, 'QN', 990-1096, 'P', 1098-1152, 'AT', 1155, 'S', 1157-1160  
A:Experimental source: liver  
R.Loidl, H.R.; Brinker, J.M.; May, M.; Pihlajaniemi, T.; Morrow, S.; Rosenbloom, J.; Myerowitz, R.  
Nucleic Acids Res. 12, 9383-9394, 1984  
A:Title: Molecular cloning and carboxyl-1-propeptide analysis of human type III procollagen  
A:Reference number: A93551; MUID:55087944; PMID:6096827  
A:Accession: A93551  
A:Molecule type: mRNA  
A:Residues: 1065-1155, 'P', 1157-1466 <LOI>  
A:Cross-references: EMBL:X01655; EMBL:X01742; NID:929584; PIDN:CAA25821.1  
R.Miskuln, M.; Dalgleish, R.; Kluge-Beckerman, B.; Rennard, S.I.; Tolstoshev, P.; Brantley, S.E.  
Biochemistry 25, 1408-1413, 1986  
A:Title: Human type III collagen gene expression is coordinately modulated with the type I collagen gene  
A:Reference number: 152393; MUID:86187804; PMID:3754462  
A:Accession: 152393

A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1161-1200 <MIS>  
A:Cross-references: GB:M1316; NID:g180415; PIDN:AAA52003.1; PID:g180416  
R:Emanuel, B.S.; Cannizzaro, L.A.; Seyer, J.M.; Myers, J.C.  
Proc. Natl. Acad. Sci. U.S.A. 92, 3385-3389, 1995  
A>Title: Human alpha 1(III) and alpha 2(V) procollagen genes are located on the long arm  
A:Reference number: 159025; MUID:85216505; PMID:3858826  
A:Accession: I79359  
A:Status: translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1165-1196 <EWA>  
A:Cross-references: GB:M1134; NID:g180417; PIDN:AAA52004.1; PID:g180418  
R:Chu, M.T.; Weil, D.; de Wac, W.; Bernard, M.; Sippola, M.; Ramirez, F.  
J. Biol. Chem. 260, 4357-4363, 1985  
A>Title: Isolation of cDNA and genomic clones encoding human pro-alpha(III) collagen.  
A:Reference number: A92516; MUID:85157600; PMID:2579949  
A:Accession: A92516  
A:Molecule type: DNA  
A:Residues: 1176-1240 /V/, 1242-1356, /P/, 1358-1466 <CHU>  
A:Cross-references: GB:M10615; GB:M10793; GB:M10794; GB:M10795; GB:M10796; GB:M10797; GB:  
A:Experimental source: liver  
A:Note: the authors translated the codon TTC for residue 1057 as Tyr; the codons given  
action  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
3-hydroxylated. About 15% of the lysines are 5-hydroxylated and some are subsequently  
C:Genetics:  
A:Gene: GDB:COL3A1  
A:Cross-references: GDB:118729; OMIM:120180  
A:Map position: 2q31-q31  
A:Introns: 27/1; 94/3; 111/3; 149/3; 176/3; 554/3; 587/3; 1175/3; 1373/3; 1418/  
A:Notes: The list of introns is incomplete; defects in this gene can result in Ehlers-Da  
C:Complex: type III collagen is a homotrimer of monomers initially linked by disulfide  
er of their length, is formed with desmosine cross-links made from lysine and allylsine  
C:Function:  
A>Description: structural component of extracellular fibrous polymer that maintains int  
C:Superfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology  
C:Keywords: coiled coil; Ehlers-Danlos syndrome; extracellular matrix; glycoprotein; hy  
R:1-23/Domain: signal sequence #status predicted <SIG>  
R:124-153/Domain: amino-terminal propeptide #status predicted <PRO>  
R:154-91/Domain: von Willebrand factor type C repeat homology <WMC>  
R:154-1221/Product: collagen alpha 1(III) chain #status predicted <MAT>  
R:154-167/Region: amino-terminal nonhelical telopeptide  
R:168-1196/Region: helical  
R:1091-1093/Region: cell attachment (R-G-D) motif  
R:1197-1221/Region: carboxyl-terminal nonhelical telopeptide  
R:1122-1466/Domain: carboxyl-terminal propeptide #status predicted <CPR>  
R:1248-1466/Domain: fibrillar collagen carboxyl-terminal homology <FCC>  
R:154-154/Cleavage site: pyroglutamate carboxylic acid (Gln) (in mature form) #status predi  
R:155/Modified site: Pro-Gln (procollagen N-endopeptidase) #status predicted  
R:161-1212/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predi  
R:163,284,960,977,1106/Modified site: allylsine (Lys) #status predicted  
R:263/Binding site: carbohydrate (Lys) #status predicted  
R:263/Binding site: carbohydrate (Lys) (covalent) #status experimental  
R:158,1094/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental  
R:948-949/Cleavage site: Gly-Ile (collagenase) #status experimental  
R:1106/Binding site: carbohydrate (Lys) (covalent) #status predicted

Query Match	Best Local Similarity	Score	Length
Matched 47;	Conservative 18;	Mismatches 51;	Indels 60;
Gaps 9;			
5	AELIGOTSRAGLCRP-----LCL-----LLCASQ---LPRTMHPOGR--NASPOLL 45	8.5%;	Score 125; DB 1; Length 1466;
Db	SHLDGSYADNDVKWPBECQICVCDSGVSLCDIICDDQLDPCNPPIPLPBECACAVCQ-- 90	26.7%;	Pred. No. 0.007;
Oy	LGLFLVLLLLLIQSAPSSASSENPKVKOKALIRREVVDLVNGKCLQGPAVGPGSDSGA 105		
Db	91 -----PTATPRPP-----NGQPQGPCKMDPDPGGPIGR 119		
Oy	106 NGIPGTGTCIGRGDFKFKGEKCLRSEF---ESSTPTKYKCSCMSNLVGIIDLKIA 158		
Db	120 NGDPGIRIGQGPSPPSPGPIC--ESCPTGPONNVSPDYSDVYS---GAVALGIA 170		

```
RESULT 5
T15268
hypothetical protein F59E12.12 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T15268
R:Johnson, D.
Submitted to: the EMBL Data Library, May 1997
A:Description: The sequence of C. elegans cosmid F59E12.
A:Reference number: Z18318
A:Accession: T15268
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-311 <OH>
A:Cross-references: UNIPROT:O01904; EMBL:AF003386; NID:92088833; PID:92088834; PIDN:AA85
A:Experimental source: strain Bristol N2; clone F59E12
C:Genetics:
A:Gene: CESP:F59E12.12
A:Map position: 2
A:Introns: 24/2

Query Match      8.4%; Score 123.5; DB 2; Length 311;
Best Local Similarity 53.3%; Pred. No. 0.0016;
Matches 24; Conservative 4; Mismatches 12; Indels 5; Gaps 1;

QY 87 GNC-----IQGPAGVGRDSSPGANGIPGTGIPGDFKGEKGE 126
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 227 GKCDEVNVVAQGPFGSGPPGLPGPDGLPGTGNPGQDGRGPAGE 271

RESULT 6
T29731
hypothetical protein F41F3.4 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T29731
R:Badcock, K.; Le, T.T.; Gattung, S.
Submitted to: the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid F41F3.
A:Reference number: Z20674
A:Accession: T29731
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-310 <BAD>
A:Cross-references: UNIPROT:Q20822; EMBL:U55366; PIDN:AAA97982.1; GSPDB:GN00023; CESP:F4
A:Experimental source: strain Bristol N2; clone F41F3
C:Genetics:
A:Gene: CESP:F41F3.4
A:Map position: 5
A:Introns: 27/3; 68/3

Query Match      8.3%; Score 122.5; DB 2; Length 310;
Best Local Similarity 24.4%; Pred. No. 0.002;
Matches 39; Conservative 9; Mismatches 39; Indels 73; Gaps 5;

QY 4 AAEIGTTSBAICRPLCLILCASQIPHYMPCGRASPOLILGLFLVLLILLISAPSS 63
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 93 AVEVASRVSVCSCQCCCL-----PGPGTGPAPGRGPKR-----GRPGA 132

QY 64 ASEN-----PKYKQKALIRQREVVVLVNGMCLQGPAGVPGRDSPGANG 107
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 133 CGNHGPGKPTGLPCDPVTVTPCKP-----CPPGAFGHGAGAGAG 176

QY 108 IPGTGPI-----DGRDFFKGEKGE 126
      |||||:|||||:|||||:|||||:|||||:|||||:
DB 177 KEGAPGVGGSSGAPGAPKAPGAPGPGQPGRDGQPGQAQ 216

RESULT 7
A45748
collagen alpha 1(VII) chain - mouse (fragment)
```

```
C:Species: Mus musculus (house mouse)
C:Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
C:Accession: A45748
R:Li, K.; Christiano, A.M.; Copeland, N.G.; Gilbert, D.J.; Chu, M.L.; Jenkins, N.A.; Ut
Genomics 16, 733-739, 1993
A:Title: cDNA cloning and chromosomal mapping of the mouse type VII collagen gene (Col7a
A:Reference number: A45748; MUID:93315168; PMID:8325648
A:Accession: A45748
A:Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-920 <LII>
A:Cross-references: UNIPROT:O63870; GB:S63654; NID:9386656; PIDN:AA827492.1; PID:9386657
A:Experimental source: epidermal keratinocyte
A>Note: sequence extracted from NCBI backbone (NCBI:135000, NCBI:P:135001)

Query Match      8.3%; Score 122.5; DB 2; Length 920;
Best Local Similarity 63.2%; Pred. No. 0.0069;
Matches 24; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 92 GPAGVGR---DGSRGANGIPGTGIPGRDFFKGEKGE 126
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 680 GPPGVPGKAGDGPGLNGKNGDPEDPBGDKRGEKGD 717

RESULT 8
I48103
type VII collagen - Chinese hamster (fragment)
C:Species: Cricetus griseus (Chinese hamster)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I48103
R:Greenspan, D.S.
Hum. Mol. Genet. 2, 273-278, 1993
A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
A:Reference number: I48103; MUID:93271985; PMID:8499916
A:Accession: I48103
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1549 <RBS>
A:Cross-references: UNIPROT:O60444; GB:I06863; NID:9386624; PIDN:AAA36968.1; PID:9386625
F:I484-1336/Domain: animal Kunitz-type proteinase inhibitor homology <BRI>

Query Match      8.3%; Score 122.5; DB 2; Length 1549;
Best Local Similarity 63.2%; Pred. No. 0.012;
Matches 24; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 92 GPAGVGR---DGSRGANGIPGTGIPGRDFFKGEKGE 126
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 426 GPPGVPGKGTGEGKPGKNGKNGEPDGPBGDKRGEKGD 463

RESULT 9
A39762
collagen alpha 1(XIV) chain - bovine (fragments)
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Feb-1992 #sequence_revision 05-Apr-1995 #text_change 09-Jul-2004
C:Accession: A39762; S23815; S16905
R:Hubert-Foucher, B.; Font, B.; Eichenberger, D.; Goldschmidt, D.; Lethias, C.; van der
J. Biol. Chem. 267, 15759-15764, 1992
A:Title: Purification and characterization of native type XIV collagen.
A:Reference number: S23815; MUID:92348439; PMID:1322405
A:Accession: S23815
A:Molecule type: protein
A:Residues: 1-19;20-61;62-71;72-81;82-123;124-178 <UB>
A:Cross-references: UNIPROT:Q7M2M9
R:Hubert-Foucher, B.; Font, B.; Eichenberger, D.; Goldschmidt, D.; Lethias, C.; van der
J. Biol. Chem. 267, 15759-15764, 1992
A:Title: Purification and characterization of native type XIV collagen.
A:Reference number: S23815; MUID:92348439; PMID:1322405
A:Accession: S23815
A:Molecule type: protein
A:Residues: 'XR',82-94,'X',96 <AUB>
C:Complex: homotrimer
C:Keywords: coiled coil; disulfide bond; extracellular matrix; homotrimer; hydroxylysine
```

```

Query Match      8.3%; Score 122; DB 2; Length 178;
Best Local Similarity 32.7%; Pred. No. 0.0012;
Matches 32; Conservative 8; Mismatches 34; Indels 24; Gaps 4;

Qy      32  TMAPGPRAPSPOLLGLFLVLLLLQLGAPSSASSENPKYKALIRQREVDVLYNGMCQ 91
      46  SWMGPGALGPP-----GVPGAGD-----LQSQAMRQ-VCEQLQSHHTQ 84
      46  SWMGPGALGPP-----GVPGAGD-----LQSQAMRQ-VCEQLQSHHTQ 84

Db      92  GPAGVPGRDSPGA--NGIPGTGIPGRDGFKEKGE 126
      85  GPGEPRPGSPGTPRGEGPPGAGFPNGAGVPGSGXGE 122

RESULT 10
54657
collagen alpha 1(XIV) chain - human (fragments)
C:Species: Homo sapiens (man)
C:Date: 15-Jul-1995 #sequence revision 03-Nov-1995 #text_change 09-Jul-2004
C:Accession: S4657; S37750; A44398
R:Brown, J.C.; Golbik, R.; Mann, K.; Timpl, R.
Matrix Biol. 14, 287-295, 1994
Article: Structure and stability of the triple-helical domains of human collagen XIV.
A:Reference number: S4657; MUID:95128518; PMID:7827751
A:Accession: S4657
A:Status: preliminary
A:Molecule type: protein
A:Residues: 26-330 <BRO>
A:Cross-references: UNIPROT:Q9UDP6
R:Brown, J.C.; Golbik, R.; Mann, K.; Timpl, R.
submitted to the EMBL Data Library, October 1993
A:Description: Structure and stability of the triple helical domains of human collagen X
A:Reference number: S37749
A:Accession: S37750
A:Molecule type: protein
A:Residues: 203-330 <BR2>
R:Brown, J.C.; Mann, K.; Wiedemann, H.; Timpl, R.
J. Cell Biol. 120, 557-567, 1993
A>Title: Structure and binding properties of collagen type XIV isolated from human placenta
A:Reference number: A44398; MUID:93132085; PMID:8421066
A:Accession: A44398
A:Molecule type: protein
A:Residues: 1-17,18-25,26-42,'X',44-45,74-84,100-117,136-150,'SM',238-260,'X',262-263,'X'
A:Experimental source: placenta
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C:Genetics:
A:Gene: COL1A1
A:Complex: type XIV collagen may be a heterotrimer of alpha 1(XIV) chains and one alpha
C:Function:
A:Description: structural component of extracellular fibrous polymer associated with col
C:Keywords: coiled coil, extracellular matrix, glycoprotein, hydroxylysine, hydroxyprol
F:374,84,111,117,141,144,147,219,222,225,228,231,237,240,243,249,252,267,278,292,295,298,
F:374,206,289,Modified site: 4-hydroxyproline (Pro) (partial) #status experimental
F:43,261,Modified site: 5-hydroxylysine (Lys) #status experimental
F:43,261,Modified site: carboxydrate (Lys) (covalent) #status experimental
F:264/Modified site: 5-hydroxylysine (Lys) (partial) #status experimental

Query Match      8.3%; Score 122; DB 2; Length 330;
Best Local Similarity 31.9%; Pred. No. 0.0024;
Matches 36; Conservative 10; Mismatches 33; Indels 34; Gaps 5;

Qy      35  PGRAPSPOLLGLFLVLLLLQLGAPSSASSENPKYKALIR--GREYVD-----L 84
      155  PGRALGPPV-----PGAKGERGERDGLQAMVXARQVXBOLLQSHMAR 201
      155  PGRALGPPV-----PGAKGERGERDGLQAMVXARQVXBOLLQSHMAR 201

Db      85  YNGM-----CLQGPAGVPGRDSPGA--NGIPGTGIPGRDGFKEKGE 126
      202  YNGIPSHSSSIRTVQGPGRGPRGSPGAPGEGQPPGTGPPGNGVGVPTPEB 254
      202  YNGIPSHSSSIRTVQGPGRGPRGSPGAPGEGQPPGTGPPGNGVGVPTPEB 254

RESULT 11
516366

```

```

collagen alpha (IV) chain precursor - pig roundworm
C:Species: Aecaris suum (pig roundworm)
C:Date: 04-Dec-1992 #sequence_rev1stion 04-Dec-1992 #text_change 09-Jul-2004
C:Accession: S16366
R:Petit, J.; Kingston, I.B.
J:Biol. Chem. 266, 16149-16156, 1991
A:Title: The complete primary structure of a nematode alpha-2(IV) collagen and the part
A:Reference number: S16366; MUID:91340768; PMID:11714907
A:Accession: S16366
A:Molecule type: mRNA
A:Residues: 1-1763 <JBI>
A:Cross-references: UNIPROT:P27393; GB:M67507; NID:G159648; PIDN:AAA18014.1; PID:G15964
C:Genes:
A:Introns: 229/3; 266/3; 305/3; 360/3; 424/1; 489/1; 548/1; 656/3; 790/1; 891/1; 963/1;
C:Superfamily: collagen alpha 1(IV) chain
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; disulf
F:1-26/Domian: signal sequence #status predicted <SIG>
F:27-1763/Product: collagen alpha 2(IV) chain #status predicted <MAT>
F:27-42/Domian: non-collagenous NH1 #status predicted <NH1>
F:43-1529/Domian: collagenous #status predicted <COL>
F:197-199/Region: cell attachment (R-G-D) motif
F:1530-1763/Domian: carboxyl-terminal nonhelical, NC1 #status predicted <NC1>
F:1530-1638/Domian: repeat NCI #status predicted <NC12>
F:1639-1763/Domian: repeat NCI #status predicted <NC12>
F:31,34,39,41,536,539/Disulfide bonds: interchain #status predicted
F:1126/Binding site: carbonylde (Asn) (covalent) #status predicted
F:1593-1599,1702-1709/Disulfide bonds: #status predicted

Query Match 8.2%; Score 121; DB 2; Length 1763;
Best Local Similarity 62.9%; Pred. No. 0.02; Mismatches 11; Indels 0; Gaps 0;
Matches 22; Conservative 2;

Qy 92 GPAGVPGRDGSPGANGIPGTPGIPGRDGFKEKGE 126
Db 162 GVAGAPGEKGDAGMGPMGLPGRDGFPEKGE 196

RESULT 12
S29328
complement subcomponent C1q chain C - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S29328
R:Petty, F.; Reid, K.B.M.; Loos, M.
Eur. J. Biochem. 209, 129-134, 1992
A:Title: Isolation, sequence analysis and characterization of cDNA clones coding for th-
ecerebellin.
A:Reference number: S29328; MUID:93011118; PMID:1396621
A:Accession: S29328
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-246 <PBT>
A:Cross-references: UNIPROT:002105; EMBL:X66295; NID:G50228; PIDN:CAA46993.1; PID:G50229
C:Superfamily: complement subcomponent C1q chain A; complement C1q carboxyl-terminal hom-
F:112-245/Domian: complement C1q carboxyl-terminal homology <C1Q>

Query Match 8.1%; Score 119.5; DB 2; Length 246;
Best Local Similarity 36.6%; Pred. No. 0.0023; Mismatches 32; Indels 21; Gaps 3;
Matches 34; Conservative 6;

Qy 37 GRAASPOLLLGLFLVLLLLDLSAPSSAENPKVKKALLRREVVLDLYNMGCLQSPAGV 96
Db 4 GPSCPO--CGUCLLLFLALPLRSQAS-----AGCTGIPGMFGA 45

Qy 97 PGRDGS---PGANGIPGTPGIPGRDGFKEKGE 126
Db 46 PGKGDHGLGSPKGEPRGIPAVPGTQGRPKGKGE 78

RESULT 13
A61396
collagen alpha 1(II) chain - golden hamster (fragments)
C:Species: Mesocricetus auratus (golden hamster)

```



CDate: 09-Sep-1994 #sequence\_revision 09-Sep-1994 #text\_change 09-Jul-2004

CAccession: A61396  
CRelease: P.E.; Hosoi, J.; Montgomery, J.C.; Wiseman, R.W.; Barrett, J.C.  
Mol. Celling. 4, 14-24, 1991

A>Title: Loss of a tumor suppressor gene function is correlated with downregulation of c

A.Reference number: A61396; MUID:91182265; PMID:2009131  
A.Accession: A61396

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A.Residues: 1-323 <CI>

A.Cross-references: UNIPROT:Q7W099

CSuperfamily: collagen alpha 1(I) chain; fibrillar collagen carboxyl-terminal homology /  
F.i20-323/Domain: fibrillar collagen carboxyl-terminal homology <FCC>

Query Match            8.1%   Score 119.5; DB 2; Length 323;  
Best local similarity   20.6%; Pred. No. 0.004; Indels 153; Gaps 10,  
Matches   57; Conservative   20; Mismatches   47;

Oy         90 LOGPAGVGRDGGSPGAN-----GIPTGP-----GIRSEFE      112  
        ||||| |::|:  
Db         9 LTGPAGEGRREGSGPADGRPGRDGAAGVGKRGRTGTALGARPAAPPGAPAGATKKQG      68  
Oy         113 -----GTPRDGFKGKEGE-----CLRSFE      133  
        ||||| :|:  
Db         69 DRGESAQGWGPSGPPAGARGLPGQPGRGDKGSGBEGSLRQHDVEVDATKLINQIE      128  
Oy         134 ESWTP-----NYKC--SMSLNVGIDIGTAECTFTFKMS-----      167  
        :|:                :|:                :|:  
Db         129 SIRSDGSCNRPARTCPDLKCHPEPKSGDWIDNQ--GCTLDAMKAFCMETGETCYV      186  
Oy         168 NSA-----LRVLF--SSGLRLKRNACCORRYFTFGNAEC      200  
        :|:                :|:                :|:  
Db         187 PNPNASVPKNMWSSKGEEKHIVQMTFLRLSTEGSQNIYTHCKN-----      231  
Oy         201 SGPLPEAIIVLDGSPELNSTINHTSVEGLECG      237  
        :|:                :|:                :|:  
Db         232 -----SIAYIDEANGLKKALLIQSSNDVMRAEG      261

RESULT 14

CGHU3B

collagen alpha 3(IV) chain precursor, long splice form - human

N.Alternate names: Goodpasture antigen, procollagen alpha 3(IV) chain long splice form

C.Species: Homo sapiens (man)

C.Date: 28-Oct-1994 #sequence\_revision 03-Oct-1995 #text\_change 09-Jul-2004

R.Accession: A54763; A43928; A44043; A45971; A439786

R.Mariyama, M.; Lehtonen, A.; Mochizuki, T.; Tryggvason, K.; Reeder, S.T.

J. Biol. Chem. 269, 23013-23017, 1994

A.Title: Complete primary structure of the human alphas3(IV) collagen chain. Coexpression

A.Reference number: A54763; MUID:94364954; PMID:8083201

A.Accession: A54763

A:Molecule type: mRNA

A.Residues: 1-1670 <MR>

A.Cross-references: UNIPROT:Q01955; GB:X80031; NID:g577563; PID:g577564

A.Experimental source: kidney

R.Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.

J. Clin. Invest. 89, 592-601, 1992

A.Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha

A.Reference number: A43928; MUID:92147878; PMID:1737849

A.Accession: A43928

A:Molecule type: mRNA

A.Residues: 1331-1524, 'I',1526-1670 <TUR>

A.Cross-references: GB:M81379

A.Experimental source: kidney

R.Ouinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saas, J.

J. Biol. Chem. 267, 19780-19784, 1992

A>Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture

A.Accession number: A44043; MUID:93015826; PMID:1400291

A:Molecule type: DNA, mRNA

A.Residues: 1386-1670 <QU>

A.Cross-references: GB:M92993; NID:g177895; PID:NAA21610.1; PID:g177896

A>Note: sequence extracted from NCBI backbone (NCBIP:115597)  
R.Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.  
J. Biol. Chem. 269, 17358, 1994  
A:Reference number: A44738; MUID:94274734; PMID:8006044  
A:Contents: annotation; extralum; correction to intronic sequence in A44043  
R.Bernal, D.; Quinones, S.; Saus, J.  
J. Biol. Chem. 269, 12090-12094, 1993  
A>Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.  
A:Reference number: A45971; MUID:93280184; PMID:8505332  
A:Accession: A45971  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1427-1444 <BBR>  
A>Note: sequence extracted from NCBI backbone (NCBIP:133363); sequence incorrectly ident  
R.Morrison, K.E.; Matiyama, M.; Yang-Feng, T.L.; Reeder, S.T.  
Am. J. Hum. Genet. 49, 545-554, 1991  
A>Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of  
A:Reference number: A39786; MUID:91353570; PMID:1882840  
A:Accession: A39786  
A:Molecule type: mRNA  
A:Residues: 1453-1593, 'A', 1595-1670 <MOR>  
A:Cross-references: GB:555790; NID:g234418; PID:AAB19637.1; PID:g234419  
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit (  
ed and subsequently O-glycosylated.  
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope  
C:Genetics:  
A:Gene: GDB:COL4A3  
A:Cross-references: GDB:128351; OMIM:120070  
A:Map position: 2q36-q37  
A:Introns: 1385/1, 1418/1, 1486/1, 1547/2, 1585/3, 1632/2 #status incomplete  
A>Note: The alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands with  
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha 3(  
mong trimer amino-terminal domains (with disulfide and desmosine cross-links), dimeric a  
er associations in the interrupted helical domain (with disulfide and desmosine cross-lin  
C:Function:  
A:Description: minor structural component of extracellular basement membrane in kidney g  
C:Superfamily: collagen alpha 1(IV) chain  
C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extracel  
F:1-18/Domain: signal sequence #status predicted <SIG>  
F:29-1670/Product: collagen alpha 3(IV) chain, long splice form #status predicted <MAT>  
F:29-42/Domain: amino-terminal nonhelical, NH1 <NH1>  
F:43-1438/Region: interrupted helical  
F:791-793/Region: cell attachment (R-G-D) motif  
F:996-998/Region: cell attachment (R-G-D) motif  
F:1154-1156/Region: cell attachment (R-G-D) motif  
F:1306-1308/Region: cell attachment (R-G-D) motif  
F:1345-1347/Region: cell attachment (R-G-D) motif  
F:1432-1434/Region: cell attachment (R-G-D) motif  
F:1439-1670/Domain: carboxyl-terminal nonhelical, NCI <NCI>  
F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:1561-1665/Domain: collagen IV carboxyl-terminal repeat <CTR>  
F:31,33,35,41,125,422,476,479,682,722,809,1387/Disulfide bonds: interchain #status predi  
F:125/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted  
F:1505-1511,1616-1622/Disulfide bonds: #status predicted  
F:1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

Query Match	Similarity	Score	DB 1;	Length
Best Local	57.1%;	Pred. NO. 0.025;		1670;
Matches	24; Conservative	3; Mismatches	12;	Indels 3; Gaps 1;

91 QGPAGVPGRDGSPGANGIPGTGPIRGD---GEKGEKSECTLR 129  
||| ||| ||| ||| : ||| ||| ||| |||  
Db 504 QGANGLKSGPSGSFGNTGLPFPGFPGAGSDPLKKGKEGRITLQ 545

RESULT 15  
AS4849  
collagen alpha 1(VII) chain precursor - human  
N:Alternate names: procollagen alpha 1(VII) chain  
C:Species: Homo sapiens (man)  
C:Date: 04-Nov-1994 #sequence\_revision 04-Nov-1994 #text\_change 09-Jul-2004  
C:Accession: A54849; PI0844; SI6316; I56328;AA0296; I84686



R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.  
 J. Biol. Chem. 269, 20256-20262, 1994  
 A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1(VII)  
 A:Reference number: A54849; PMID:94327588; PMID:8051117  
 A:Accession: A54849  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-2944 <CHR>  
 A:Cross-references: UNIPROT:Q02388; GB:L02870; NID:g987124; PIDN:AAA75438.1; PID:g987125  
 R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Imanura, S.  
 Biochem. Biophys. Res. Commun. 183, 958-963, 1992  
 A:Title: Molecular cloning and characterization of type VII collagen cDNA.  
 A:Reference number: PH0844; PMID:92231902; PMID:1567409  
 A:Accession: PH0844  
 A:Molecule type: mRNA  
 A:Residues: EFR, 340-475, RAUSTASHSLCTMRATRMHPGNSHWTACPCNPASHRAARAG, 524-528, 'C',  
 A:Cross-references: DDBJ:D11152; DDBJ:D13694; NID:g453698; PIDN:BA02853.1; PID:g453699  
 A:Experimental source: Keratinocyte  
 A:Note: the authors translated the codon ACC for residues 394 and 397 as Tyr  
 R:Parente, M.G.; Chung, L.C.; Rymaszewski, J.; Woodley, D.T.; Wynn, K.C.; Bauer, E.A.; Mat  
 Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991  
 A:Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.  
 A:Reference number: S16316; PMID:91334380; PMID:1871109  
 A:Accession: S16316  
 A:Molecule type: mRNA  
 A:Residues: 815-892, 'E', 894-1439 <PAR>  
 A:Cross-references: GB:M65158; GB:S45017; NID:g180914; PIDN:AAA96439.1; PID:g180915  
 A:Experimental source: Keratinocyte  
 R:Gammot, W.R.; Abernethy, M.L.; Padilla, K.M.; Prissayanh, P.S.; Cook, M.E.; Wright, J.;  
 J. Invest. Dermatol. 99, 691-696, 1992  
 A:Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot  
 A:Reference number: I56328; PMID:93107742; PMID:1469284  
 A:Accession: I56328  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 'EFR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>  
 A:Cross-references: GB:S51236; NID:g262308; PIDN:AA2637.1; PID:g262309  
 R:Selzer, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgess, R.E.  
 J. Biol. Chem. 264, 3822-3826, 1989  
 A:Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagenase  
 A:Reference number: A30296; PMID:8919437; PMID:2537292  
 A:Accession: A30296  
 A:Molecule type: protein  
 A:Residues: 'A', 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 2032, 'C', 2034-2041, '  
 A:Note: two reported peptides cannot be reliably located  
 R:Greenspan, D.S.  
 Hum. Mol. Genet. 2, 273-278, 1993  
 A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous  
 A:Reference number: I48103; PMID:93271985; PMID:8499916  
 A:Accession: I48103  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 2395-2871, 'S', 2873-2944 <RE2>  
 A:Cross-references: GB:L06862; NID:g388713; PIDN:AAA69196.1; PID:g388714  
 R:Christiano, A.M.; Rymaszewski, M.; Uitto, J.  
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994  
 A:Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs  
 A:Reference number: A55255; PMID:94224777; PMID:8170945  
 A:Contents: annotation  
 A:Comment: Prolines and lysines at the third position of the tripeptide repeating unit  
 ed and subsequently O-glycosylated.  
 C:Genetics:  
 A:Gene: GDB:COL7A1; EBR1; EBD1; EB  
 A:Cross-references: GDB:128750; OMIM:120120  
 A:Map position: 3p21.3-3p21.3  
 A:Note: defects in this gene can result in dominant and recessive dystrophic epidermolys  
 A:Note: there are 118 introns  
 C:Complex: type VII collagen is probably a homotrimer  
 C:Function:  
 A:Description: structural component of extracellular polymer associated with anchoring E  
 C:Keywords: coiled coil; extracellular matrix; glycoprotein; hydroxylysine; hydroxyprolin  
 F:1-16/Domain: signal sequence #status predicted <SIG>  
 F:17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>

F:17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>  
 F:36-201/Domain: von Willebrand factor type A repeat homology <VWA1>  
 F:321-318/Domain: fibronectin type III repeat homology <FN1>  
 F:327-413/Domain: fibronectin type III repeat homology <FN2>  
 F:414-502/Domain: fibronectin type III repeat homology <FN3>  
 F:508-593/Domain: fibronectin type III repeat homology <FN4>  
 F:598-683/Domain: fibronectin type III repeat homology <FN5>  
 F:686-771/Domain: fibronectin type III repeat homology <FN6>  
 F:776-862/Domain: fibronectin type III repeat homology <FN7>  
 F:864-952/Domain: fibronectin type III repeat homology <FN8>  
 F:954-1045/Domain: fibronectin type III repeat homology <FN9>  
 F:1052-1219/Domain: von Willebrand factor type A repeat homology <VWA2>  
 F:1170-1172/Domain: cell attachment (R-G-D) motif  
 F:1189-1253/Domain: cysteine/proline-rich  
 F:1254-2783/Domain: interrupted helical  
 F:1334-1336/Region: cell attachment (R-G-D) motif  
 F:2008-2010/Region: cell attachment (R-G-D) motif  
 F:2553-2555/Region: cell attachment (R-G-D) motif  
 F:2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>  
 F:2876-2929/Domain: animal Kunitz-type proteinase inhibitor homology <BPI>  
 F:337-786/1109/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:2167,2176,2185,2188,2664,2667,2673/Modified site: 4-hydroxyproline (Pro) #status expe  
 F:2625,2631/Modified site: 5-hydroxylysine (Lys) #status experimental  
 F:2625,2631/Binding site: carbohydrate (Lys) (covalent) #status experimental  
 F:2634,2802,2804/Disulfide bonds: interchain #status predicted

Query Match 8.1%; Score 119.5; DB 2; Length 2944;  
 Best Local Similarity 60.5%; Pred. No. 0.048;  
 Matches 23; Conservative 3; Mismatches 9; Indels 3; Gaps 1;  
 QY 92 GPAGVPR---DGPAGANGPCTPGPGDGRKGE 126  
 DB 1824 GPAGVPR---DGPAGANGPCTPGPGDGRKGE 1861

Search completed: August 4, 2005, 04:22:30  
 Job time : 23.0876 secs

This Page Blank (uspto)



09D1D6 PRELIMINARY; PRT; 245 AA.  
 ID 09D1D6  
 AC 09D1D6;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE Mus musculus 18-day embryo whole body cDNA, RIKEN full-length enriched  
 DE library, clone:1110014B07 product:hypothetical collagen triple helix  
 DE repeat containing protein, full insert sequence.  
 GN Name=Ctrnc1;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning.";  
 RL Meth. Enzymol. 303:19-44(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;  
 RA RIKEN FANTOM Consortium;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA The FANTOM Consortium.  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 RL Nature 420:563-573(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;  
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new genes.";  
 RL Genome Res. 10:1617-1630(2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;  
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,  
 RA Konno H., Akiyama J., Nishi K., Kitsuina T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Hatada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsumura S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 sequencing pipeline with 384 multicapillary sequencer.";  
 RL Genome Res. 10:1757-1771(2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Whole body;  
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,  
 RA Arawaka T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kashiwagi T., Kato H.,  
 RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,  
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,  
 RA Sano H., Saeki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,  
 RA Sogade Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,  
 RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,  
 RA Muramatsu M., Hayashizaki Y.;

RL Submitted (JUL-2000) to the EMBL/GenBank/DBSJ databases.  
 DR EMBL; AK003674; BAB22930.1; --  
 DR WGI; MG1:1915838; Ctrnc1.  
 DR GO; GO:0005737; Cytochrome, IEA.  
 DR GO; GO:0006817; P; phosphate transport; IEA.  
 DR InterPro; IPR008161; Clg helix.  
 DR InterPro; IPR008160; Collagen.  
 DR Pfam; PF01391; Collagen: 1.  
 DR ProDom; PD000007; Clg\_helix; 1.  
 KW Collagen, Hypothetical protein.  
 SQ SEQUENCE 245 AA; 26460 MW; 14951B87D8181A0E CRC64;  
 Query Match 98.9%; Score 1292; D:2; Length 245;  
 Best Local Similarity 98.8%; Pred. No. 9.9e-107;  
 Matches 242; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 MHPOGRAAPQQLLLGLFVLLLLQLSAPSSASNPVKYKQALIRQREVLDYNGMCLQG 60  
 DB 1 MHPOGRAAPQQLLLGLFVLLLLQLSAPISASNPVKYKQALLIRQREVLDYNGMCLQG 60  
 QY 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGRKGEKGECLRESFESWTPTNTKQCSWSLNYGI 120  
 DB 61 PAGVPGRDGSPGANGIPGTPTGIPGRDGRKGEKGECLRESFESWTPTNTKQCSWSLNYGI 120  
 QY 121 DLGKTAECTFTMRNSALRVLFSGSLRLPKRCNACCQWYFFNGACSGPLPIAITYL 180  
 DB 121 DLGKTAECTFTMRNSALRVLFSGSLRLPKRCNACCQWYFFNGACSGPLPIAITYL 180  
 QY 181 DGSGPELNSTINIRHTSSVEGLCEGIGALVDVAIWTGCDSPKGDASTGNSVSRILI 240  
 DB 181 DGSGPELNSTINIRHTSSVEGLCEGIGALVDVAIWTGCDSPKGDASTGNSVSRILI 240  
 QY 241 EELPK 245  
 DB 241 EELPK 245  
 RESULT 3  
 ID 096CG8 PRELIMINARY; PRT; 243 AA.  
 AC 096CG8;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)  
 DE Collagen triple helix repeat containing 1 (collagen triple helix  
 DE repeat-containing protein 1).  
 GN Name=CTRCL;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Kidney;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnae.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
 RA Altschul S.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,  
 RA Hopkins R.F., Zeeberg B., Buettow K.H., Scheefter C.F., Bhat N.K.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ustin T.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Raha S.S., Lounellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Mizny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,  
 RA Krzywinski M.I., Skalski U., Smallus D.E., Scherch N., Scherch J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human

[illegible]

RA	"Godowski P.: RT effort to identify novel human secreted and transmembrane proteins: a RT bioinformatics assessment."
RL	Genome Res. 13:2265-2270(2003).
DR	EMBL; AY38914; AAC08273.1; "
DR	GO: 0005737; C:cytoplasm; IEA.
DR	GO: 0006817; P:phosphate transport; IEA.
DR	InterPro; IPR008161; Clg_helix.
DR	InterPro; IPR008160; Collagen.
DR	Prodrom; PD000007; Clg_helix; 1.
KW	COLLAGEN.
SO	SEQUENCE 243 AA; 26266 MW; BCB49AF4DBC303BC CRC64;
Query Match	93.0%; Score 1216; DB 2; Length 243; Best Local Similarity 94.3%; Pred. No. 5.7e-100; Matches 231; Conservative 3; Mismatches 9; Indels 2; Gaps 1;
Dn	1 MHPOGRASPOLLGLFVLLTLQLSAPSSASNPVKOKALRREVRVDLYNMCLOG 60 1 MRPGPAPSFORLNG--LIILILLIQLPAFPSASISIPKQKQOLRORERVDLYNMCLOG 58
Oy	61 PAGVGPDGSPGANIGIPETPGIPGRDGFKEGECLRESFEBSWTPNYKCSWSLNYGI 120 59 PAGVGPDGSPGANVVIPTGPIDPGRDGPKGEKGCLERESFEBSWTPNYKCSWSLNYGI 118
Dn	121 DLGLIACTFTPKMSNSNALRYLPFGSGRLPKCRNACCOQRWTFNFAGBESSGLPIEAIYYL 180 119 DLGLIACTFTPKMSNSNALRYLPFGSGRLPKCRNACCOQRWTFNFAGBESSGLPIEAIYYL 178
Oy	181 DQGSPELNSTINIRHTSVSEGLCGIGAGLVDAIVWGTCSDYFKGDASTMNSVSRIII 240 179 DQGSPELNSTINIRHTSVSEGLCGIGAGLVDAIVWGTCSDYFKGDASTMNSVSRIII 238
Dn	241 EELPK 245 239 EELPK 243
Oy	241 EELPK 245 239 EELPK 243
RESULT 5	
Q81X63	PREDIMINARY; PRT; 232 AA. Q81X63
ID	Q81X63
AC	Q81X63
DT	01-MAR-2003 (TREMBlrel_23, Created)
DT	01-MAR-2003 (TREMBlrel_23, Last sequence update)
DT	01-OCT-2003 (TREMBlrel_25, Last annotation update)
DS	NWC1.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RL	Sanuki N., Fujiki K., Kanai A., Tanaka Y., Iwata T.; Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR	EMBL; AF395488; AAC017919.1; "
DR	GO: 0005737; C:cytoplasm; IEA.
DR	GO: 0006817; P:phosphate transport; IEA.
DR	InterPro; IPR008161; Clg_helix.
DR	InterPro; IPR008160; Collagen.
DR	Pfam; PF01391; Collagen; 1.
DR	Prodrom; PD000007; Clg_helix; 1.
KW	COLLAGEN.
SO	SEQUENCE 232 AA; 25163 MW; B9D4BC30304837ED CRC64;
Query Match	81.3%; Score 1063; DB 2; Length 232; Best Local Similarity 87.2%; Pred. No. 2.2e-86; Matches 197; Conservative 7; Mismatches 12; Indels 10; Gaps 1;
Oy	29 PSSASNPVKOKALRRE-----VDLYNMCIOGPAGVGPDGSPGANIGP 78 3 PGSRITVKKRKRVKSCKLENGPSAQGLICGTYNMCIOGPAGVGPDGSPGANIGP 62





```
DR GO: GO:0005737, C:cytoplasm, IEA.
DR GO: GO:0006817, P:phosphate transport, IEA.
DR InterPro: IPR008161, Clg_helix.
DR InterPro: IPR008160, Collagen.
DR Pfam: PF01391, Collagen; 2.
DR Pfam: PF02191, OLF_1.
DR ProDom: PD000007, Clg_helix; 1.
DR SMART: SM00284, OLF_1.
DR COLLAGEN.
KW SEQUENCE 551 AA; 58957 MW; CE14A36120DEC18 CRC64;

Query Match 10.0%; Score 130.5; DB 2; Length 551;
Best Local Similarity 30.7%; Pred. No. 0.0052;
Matches 35; Conservative 18; Mismatches 20; Indels 41; Gaps 5;

QY 3 PGRASAPOLLGLFLVLLLLQLSAPSSASEN-----PKVKOK-----AL 43
DB 76 PRGASAPQ-----DPAASARKSHSGEPAPHIAESHDMLMMWTYSM 119
QY 44 IROREVDLYN--GMCLGPGAVPGRDGSFGANGIPGTPGIPGRDGFKEKGE 94
DB 120 VPIRVAVVLDLGNSTKGICLTNGPSGP--GPPGAGLPGHNGLDGPGPGKGE 170

RESULT 11
Q09250 PRELIMINARY; PRT; 1669 AA.
AC Q09250;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Alpha 3 collagen IV.
OS Name=Col4a3;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
MEDLINE=20005934; PubMed=10534397; DOI=10.1006/geno.1999.5943;
RA Lu W., Phillips C.L., Killen P.D., Hlaing T., Harrison W.R.,
RA Elder F.F.B., Miner J.H., Overbeek P.A., Meisler M.H.;
RT "Insertional mutation of the collagen genes Col4a3 and Col4a4 in a
RT mouse model of Alport syndrome.";
RL Genomics 61:113-124(1999).
DR EMBL: AF169387, AAD50449.1; -.
DR PIR: I48302, I48302.
DR HSSP: P02462, 1L1L.
DR MGD: MGI:104688; Col4a3.
DR GO: GO:0005877, C:collagen type IV, ISS.
DR GO: GO:0005615, C:extracellular space, TAS.
DR GO: GO:0005178, F:integrin binding, ISS.
DR GO: GO:0008191, F:metalloendopeptidase inhibitor activity, ISS.
DR GO: GO:0006919, P:caspase activation, ISS.
DR GO: GO:0008283, P:cell proliferation, ISS.
DR GO: GO:0007166, P:cell surface receptor linked signal transdu. . .; ISS.
DR GO: GO:0006917, P:induction of apoptosis, ISS.
DR GO: GO:0016525, P:negative regulation of angiogenesis, ISS.
DR InterPro: IPR008161, Clg_helix.
DR InterPro: IPR008160, Collagen.
DR InterPro: IPR004828, Surface_antigen.
DR InterPro: IPR001442, Procollagen_C.
DR InterPro: IPR000504, RNA_rec_mot.
DR Pfam: PF01413, C4; 2.
DR Pfam: PF01391, Collagen; 21.
DR ProDom: PD000007, Clg_helix; 9.
DR ProDom: PD153432, Surface_antigen; 2.
DR ProDom: PD003923, Procollagen_C; 2.
DR SMART: SM00111, C4; 2.
DR PROSITE: PS00030, RRM_RNP_1; UNKNOWN_1.
KW Collagen.
```

```
SQ SEQUENCE 1669 AA; 161769 MW; 30976E59739A47B2 CRC64;

Query Match 9.9%; Score 129.5; DB 2; Length 1669;
Best Local Similarity 37.0%; Pred. No. 0.022;
Matches 37; Conservative 9; Mismatches 29; Indels 25; Gaps 4;

QY 1 MHQGRASPOLLLGLFLVLLLLQLSAPSSASENPKVKOKALIROREVDLYNGMCT-- 58
DB 1 MHSK---TAPRL--VFLLTLTLLLAASPVASKGCVCVKGR-----GQCICA 42

QY 59 -----QGPAGVPGRDSPGANGIPGTPGIPGRDGFKEKGE 93
DB 43 GTKGKGEKGVGSPGPGQKGFPGPEGLPGQKKGSPG 82

RESULT 12
Q08VD7 PRELIMINARY; PRT; 289 AA.
AC Q08VD7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Mus musculus 16 days embryo lung cDNA, RIKEN full-length enriched
DE library, clone:8430425G24 product:COMPLEMENT-C1Q TUMOR NECROSIS
DE FACTOR-RELATED PROTEIN homolog.
OS Name=C1qtnf7;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carrinci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carrinci P., Shibata Y., Hayatsu N., Sugihara Y., Shibata K., Itoh M.,
RA Komoto H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carrinci P.,
RA Komoto H., Akiyama J., Nishi K., Kitsuami T., Teshiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasahara K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanabe M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
```



```

RT sequencing pipeline with 384 multicapillary sequencer."
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Lung;
RA Aachari J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,
RA Atakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ono M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shingawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Yoshida Y., Toyota T., Yamamura T., Yamataka I., Yasunishi A.,
RA Yoshida K., Yoshino M., Yamamoto M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK078818; BAC37409.1; -.
DR HSSP; O60994; IC28.
DR MGD; MGI:1925911; C1qtnf7.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR001073; C1q.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; C1q; 1.
DR Pfam; PF01391; Collagen; 2.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; C1Q; 1.
DR KX
SQ SEQUENCE 289 AA; 30483 MW; 514AC7FC318832A CRC64;

Query Match 9.8%; Score 128.5; DB 2; Length 289;
Best Local Similarity 40.3%; Pred. No. 0.0038;
Matches 31; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 18 LVTLTLTLGSGSPASSENPKVKOKALIROREVDLVNMGCLGSPAGVGRDGPNGANIP 77
DB 2 IVLLVYTSIAICSGGPRPOAKGSEISPRYICISPG--LPGPDPGANGSGPPHGR 58
QY 78 GTPGIPGRDGFKEGKE 94
DB 59 GLPGRDGRDGRKEGKE 75

RESULT 13
Q18799 ID PRELIMINARY; PRT; 289 AA.
AC Q18799;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JUN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C53B4.5.
GN ORFNames=C53B4.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OC NCB1_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Berks M.;
RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z68215; CAA92453.1; -.
DR PIR; T20177; T20177.

```

```

DR WormBase; WBGene0000693; C53B4.5.
DR WormBep; C53B4.5; CE03091.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0042302; F:structural constituent of cuticle; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002466; Col_cuticle_N.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF01484; Col_cuticle_N; 1.
DR KX
SQ SEQUENCE 289 AA; 28993 MW; 2999A3FP9CC0B044 CRC64;

Query Match 9.8%; Score 127.5; DB 2; Length 289;
Best Local Similarity 48.1%; Pred. No. 0.0047;
Matches 25; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

QY 59 OGPAGVGRDGS-----PGANGIPGTPGIPGRDGFKEGKEC 95
DB 220 RGPAGRPKDGACGGRGKANGEPQGPGRDQPGRPGRDGRGHPGKGV 271

RESULT 14
Q9NAR3 ID PRELIMINARY; PRT; 291 AA.
AC Q9NAR3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein C24F3.6.
GN ORFNames=C24F3.6;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OC NCB1_taxid=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA McMurtry A.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL022716; CAB97232.1; -.
DR WormBase; WBGene0000698; C24F3.6.
DR WormBep; C24F3.6; CE18523.
DR GO; GO:0005737; C:cytoplasm; IEA.
DR GO; GO:0042302; F:structural constituent of cuticle; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01391; Collagen; 2.
DR Pfam; PF01484; Col_cuticle_N; 1.
DR KX
SQ SEQUENCE 291 AA; 29191 MW; 8A8F3685823A3534 CRC64;

Query Match 9.6%; Score 125.5; DB 2; Length 291;
Best Local Similarity 48.1%; Pred. No. 0.0071;
Matches 25; Conservative 4; Mismatches 8; Indels 15; Gaps 1;

QY 59 OGPAGVGRDGS-----NGIPGTPGIPGRDGFKEGKEC 95
DB 222 RGPAGRPKDGSPGCGGKAGSGEPQGPGRDQPGRPGRDGRGHPGKGV 273

RESULT 15
O18286 ID PRELIMINARY; PRT; 298 AA.
AC O18286;

```

```

DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein ZK1010.7.
GN ORFNames=ZK1010.7;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RL investigating biology."
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z82083; CAB04973.1; -.
DR PIR; T27644; T27644.
DR WormBase; WBGene0000672; ZK1010.7.
DR WormPep; ZK1010.7; CE15507.
DR GO; GO:0005737; Cytoplasm; IEA.
DR GO; GO:0042302; F:structural constituent of cuticle; IEA.
DR GO; GO:0006817; P:phosphate transport; IEA.
DR InterPro; IPR008160; Collagen.
DR InterPro; IPR002486; Col_cuticle_N.
DR Pfam; PF01484; Col_cuticle_N; 1.
DR Collagen; Hypothetical protein.
KW SEQUENCE 298 AA; 30260 MW; BE5FD5B5D5C635DD CRC64;
SQ
Query Match 9.64; Score 125.5; DB 2; Length 298;
Best Local Similarity 39.54; Pred. No. 0.0073;
Matches 30; Conservative 6; Mismatches 29; Indels 11; Gaps 2;
QY 31 SASGNPKVQKALIRQEVVDLYNGMCLQGPGVPGHSGPGANGIPGTGIPGRDQFKG 90
Db 82 SSNETTRLKRDAAACQ-----GCCL---PGLPGDPGPPGXGXGAPRGAPGAPQFP 130
QY 91 EKGECLRESFESWTP 106
Db 131 RPPAVCEITEPPCTP 146

```

Search completed: August 4, 2005, 04:21:22  
 Job time : 93.3529 secs